

EXHIBIT E2

**Firearm/Toolmark Identification:
Passing the Reliability Test Under Federal and State Evidentiary Standards^a**

Authors: Richard Grzybowski^b, Jerry Miller^c, Bruce Moran^d, John Murdock^e, Ron Nichols^f, Robert Thompson^g

Keywords: Daubert, Frye, Firearm and toolmark identification, scientific method, scientific reliability, validity

ABSTRACT

Regardless of the jurisdiction in which a firearm and toolmark examiner practices, it is essential that the work product is scientifically reliable and valid. In pursuit of obtaining more reliable expert witness testimony, state and federal courts have adopted certain evidentiary standards with which all forensic disciplines must comply. This article is designed to assist examiners explain the reliability of firearm and toolmark identifications.

Introduction

This paper offers an approach to explaining the firearm and toolmark identification process as a reliable science under the challenges of both Daubert and Frye. It is important to remember that these significant court decisions were brought about by the court's zeal to ensure reliability in the methods of forensic science. Although the mechanisms of these decisions differ, the goal is the same.

The authors have no doubt that the firearm and toolmark examination discipline meets the rigorous tests of both of these court challenges. To that end, this paper provides the examiner with a discussion of practical considerations in addressing: 1) testability of the scientific principle using the scientific method, 2) known or potential error rate and the existence and maintenance of standards of control, 3) peer review and publication, and 4) general acceptance in the relevant scientific community, that is in the firearm and toolmark identification discipline. It also attempts to consolidate, under one writing, specific references in the current literature supporting the reliability of this discipline as a scientific process. These have been summarized for examiner convenience in the Appendices sections of this paper. Further, the Appendices offer some practical considerations and suggestions for preparing the examiner to provide expert testimony.

Prelude

Our profession does not exist in a vacuum. The individual examiner is bound by the judicial system of the country in which he or she practices. Therefore, like it or not, we must adhere, and often respond, to what that system dictates. The American jurisprudence is an adversarial system in that, generally, two sides present their evidence to a trier of fact (a judge or a jury) who then makes a legal finding as to the facts. Both sides have an equal right to call expert witnesses if such testimony can assist the trier of fact. This "assistance" to a jury (or a judge) is crucial in the admissibility of expert testimony. The key to

^a The views expressed in this article are those of the authors alone; no endorsement is being made by the Bureau of Alcohol, Tobacco, Firearms and Explosives or any other agency or organization.

^b Richard Grzybowski, Chief, Identification Section, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Walnut Creek, CA., California licensed attorney.

^c Jerry Miller, Firearm and Toolmark Examiner, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Atlanta, GA.

^d Bruce Moran, Criminalist IV, Sacramento County District Attorney Laboratory of Forensic Services, Sacramento, CA.

^e John Murdock, Firearm and Toolmark Examiner, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Walnut Creek, CA.

^f Ronald Nichols, Firearm and Toolmark Examiner, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Walnut Creek, CA.

^g Robert Thompson, Firearm and Toolmark Examiner, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Walnut Creek, CA.

admissibility of expert testimony is its reliability; the scientific method must produce consistent results when replicated, and the ability of an expert to explain it in detail in court.

There are several methods by which an expert's forthcoming testimony may be challenged before it is found to be admissible. The first is by a separate pre-trial evidentiary hearing (Daubert or Frye). This is a court hearing, always held outside of the jury, wherein one side (usually the defense in a criminal trial) questions the admissibility of expert testimony and requires a trial judge to make an admissibility ruling. Both sides usually present expert witnesses to prove their contention. The second is by a motion *in limine*, wherein one side moves to limit the extent of the expert's testimony. This is a more limited version of the full evidentiary hearing although it may be just as complex. It is also conducted outside of the presence of the jury. The third is by cross-examination of the expert regarding his/her qualifications and the basis for the expert's proffered testimony (*voir dire*). This is conducted via cross examination by the opposing side, in front of the jury, immediately following the side who called the expert, after they had an opportunity to present the witness' qualifications and the basis for testimony, but before the court qualifies the witness as an expert and allows the expert to present the results of examinations.

Recently the courts have taken a second look at identification sciences and rendered two opinions of some consequence to practitioners in the field of fingerprint identification and toolmark (firearms) identification. The fingerprint decision was handed down at the Federal District Court (trial court) level as the result of a Daubert evidentiary hearing. The opinion, written on January 7, 2002, by Senior Judge J. Pollak of the U.S. District Court for the Eastern District of Pennsylvania in the case of U.S. of America v. Carlos Plaza¹, *et al.*, granted, in part, the defense motion to limit testimony of fingerprint examiners to pointing out to the jury similarities and differences between the latent and known fingerprints. It specifically precluded either the prosecution or the defense experts from making any statements as to the identity of the person who left that latent print. This was a trial court decision not binding on any jurisdiction or case other than the Plaza case, but it certainly created quite a stir in the fingerprint community searching for a scientific basis for their identifications. In March 2002, as the result of a motion for reconsideration by the Government, Judge Pollak reversed himself and allowed the latent print identification in trial. However, the original ruling is significant in that it shows how a simple evidentiary hearing may affect a specific field of expertise if it appears that crucial elements of reliability are not met.

The Ramirez case² (Joseph J. Ramirez v. State of Florida, Florida Supreme Court Case No. SC92975, December 20, 2001) serves as a wakeup call to those firearm and toolmark examiners who rely on nothing more than their own subjective criteria for striae identification and are unable to put forth a convincing, logical, scientifically based explanation for the basis of their identifications. If there ever was a firm judicial rejection of the premise "I know it is a match because I have sufficient background, training and experience", the Ramirez case is it. Since this is a State Supreme Court decision, it is binding on all courts in the State of Florida (a Frye standard jurisdiction) and certainly can be cited even in those other states that adhere to the Frye evidentiary standard.

The firearm and toolmark examiner will almost certainly continue to be confronted with challenges in the current legal environment that had not been previously faced on a regular basis. The purpose of this paper is to help the individual examiner be better prepared for such confrontations. To fulfill that purpose, this paper will focus on several areas including a discussion of the legal issues, the scientific foundation of firearm and toolmark identification, error rate and standards of control, peer review and publication, and general acceptance. Further, this paper will discuss the process of describing what it is that we do as firearm and toolmark examiners. Resources and references are presented that can be used as a foundation for response. These resources are categorized and summarized in Appendix No. 1. Appendix No. 2 is a practical description of the application of the scientific method to firearm and toolmark examination.

Legal Issues

The Daubert³ and Kumho⁴ standards have already been well discussed in the literature⁵. Those jurisdictions that adhere to these standards include all Federal District Courts (trial level), Federal Circuit Courts (appellate level), Federal Military Justice Courts (following the Uniform Code of Military Justice – UCMJ), and those states that either statutorily or by case law adopted Rule 702 of the Federal Rules of Evidence (FRE). All federal courts follow the holding of Daubert/Kumho that sets out FRE 702 as the main guide for evidentiary reliability.

FRE 702 deals with testimony by experts and reads, "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, provided that (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods to the facts of the case." In addition, the trial judge may consider one or

more specific Daubert factors which are: testability of the principle, known or potential error rate and the existence and maintenance of standards of control, peer review and publication, and general acceptance in the scientific community.

Even without the Daubert factors, FRE 702 in itself requires expert testimony to be based on sufficient facts or data and be the product of reliable principles and methods. Therefore, experts must be fully prepared to provide these basic indicators of reliability in their testimony.

Two important legal factors must be mentioned. First, admissibility rulings at Daubert hearings are not binding on any parties or jurisdiction other than the parties to the present proceedings. However, in the Internet era, these decisions are immediately noticed by the media, posted on the Internet, and used by skillful lawyers to object to the admissibility of forensic identification evidence in their own cases. Second, the Kumho case made Rule 702 and the Daubert factors applicable to all expert evidence, not just scientific but also to evidence based on technical and other specialized knowledge.

It has been mistakenly believed that, under the Frye rule⁶, identification specialties (firearm/toolmark identification field among them) are untouchable because over the years so many courts have readily accepted them as complying with the Frye rule (general acceptance in the scientific community in which it belongs). However, even under the Frye rule, the parties were always free to, and sometimes did, explore the basis for an expert's opinion and question the expert's qualifications, including the particular methodology or technique used. It was usually too cumbersome, however, for a lawyer to embark on this type of a project, risking that the court would invoke the Frye rule anyway and allow identification testimony to be admitted. As mentioned earlier, this is now changing, and more lawyers are willing to expend the effort.

In the Ramirez case, the Florida Supreme Court, a Frye jurisdiction, allowed an inquiry into the basis for a positive identification of knife marks with a particular knife and, in a well written and reasoned opinion, disallowed such testimony. In Ramirez, the court was particularly troubled by the fact that the firearm examiner testified to the absolute certainty of his identification and that there are no objective criteria that must be met. Furthermore, the court found it very disconcerting that there were no photographs of the identification and that the identification was declared a "match" simply because the witness said it was a match with nothing more to support it. One of the testifying examiners stated, "it is scientific to say 'it was a match because I say so'" rather than using objective criteria and describing the basis for making an identification.

Finally, the Ramirez court found that the record did not show that the examiner's test methodology had ever been subjected to meaningful peer review or publication; a prerequisite to scientific acceptance. Reading the Florida Supreme Court opinion in the Ramirez case leaves the reader with the distinct impression that the line between the Frye and the Daubert/Kumho criteria for admissibility of expert testimony is now blurred. The wording concerning peer review and publication is taken verbatim from Daubert, and the concept of objective criteria for identification is the very basis for abandoning Frye in favor of the Federal Rules of Evidence and the Daubert factors in federal courts. The Florida decision is certainly binding on all state courts in Florida in non-firearm toolmark identification cases but it is very likely that firearm toolmark identification cases could be brought under the same umbrella.

Regardless of the jurisdiction, firearm and toolmark identification will require a scientific foundation to receive continued acceptance in the courts. The next section provides what the authors believe to be a sound scientific foundation that, if properly described, will meet and withstand any legal challenge.

Scientific Foundation of Firearms and Toolmark Identification

Prior to moving forward with foundation, it is important to define forensic toolmark identification. Forensic toolmark identification is a scientific discipline that is concerned with the identification of a toolmark to the specific tool that made it. Firearm identification is a specialized area of toolmark identification dealing with firearms, a specific category of tools. The principles and theory behind this science are the same for both areas.

Science is a formalized method of observation and experimentation that provides a theoretical explanation of what is being investigated. This formalized investigation involves application of the scientific method. The scientific method consists of: 1) stating the problem, or what it is that will be investigated; 2) gathering information about the problem; 3) forming a hypothesis that attempts to provide an explanation for the problem; 4) developing controlled experiments to test the hypothesis; 5) recording and analyzing data; 6) forming a conclusion about the validity of the hypothesis; and 7) if the hypothesis is proven false, form a new hypothesis and repeat steps 4, 5, and 6. In the event that sufficient controlled testing has been conducted that fails to disprove the hypothesis, a theory is developed that can be used to make predictions for

solving similar problems. Again, in the spirit of scientific endeavor, the validity of the theory should be subjected to continuous investigation and testing using the same steps described above, in the ongoing process of finding a better explanation to predict a solution to the original problem.

When the history of toolmark identification is examined, one will recognize that it was developed according to the precepts of the scientific method as outlined in the above discussion. The concept first hypothesized was that a toolmark could be individually identified to the specific tool by which it was produced. In other words, when we render opinions of toolmark identity, we are concluding that the tool responsible for making the mark is individualized to the exclusion of all other tools. Of course to prove this absolutely is not possible from a practical standpoint. Theoretically, to be able to make such a conclusion, the examiner would have to compare an evidence mark to all existing tools in the world. This has been recognized as impossible in writings as early as Gunther⁷ and Hatcher.⁸ Since we are not able to examine all tools in the world, we must infer an unequivocal identification. We must, therefore, be able to predict individualization without having to compare all tools that exist. Nichols^{9 10} has summarized the scientific studies that allow us, assuming no subclass influence, to predict that: 1) the working surfaces of different tools produce discernibly different toolmarks even though some quality/quantity of microscopic agreement may be present (these toolmarks are referred to as known non-matches) and; 2) toolmarks produced by the same tool working surface (referred to as known matches) can be identified with one another and exhibit a greater quality/quantity of microscopic agreement than known non-matching toolmarks.

Some of the references summarized in the appendix (and by Nichols) are examples of the extensive testing done in this area (see *Appendix No. 1, Section II.1 "Use Of The Scientific Method In Validating The Ability To Individualize Tools To The Exclusion Of All Others"*).

As a result of these studies, the Association of Firearm and Toolmark Examiners (AFTE) formulated and adopted a Theory of Identification¹¹ to explain the basic theory that allows opinions of common origin to be made in toolmark comparisons. The AFTE Theory of Identification, adopted in 1992, states:

1. The theory of identification as it pertains to the comparison of toolmarks enables opinions of common origin to be made when the unique surface contours of two toolmarks are in "sufficient agreement".
2. This "sufficient agreement" is related to the significant duplication of random toolmarks as evidenced by a pattern or combination of patterns of surface contours. Significance is determined by the comparative examination of two or more sets of surface contour patterns comprised of individual peaks, ridges and furrows. Specifically, the relative height or depth, width, curvature and spatial relationship of the individual peaks, ridges and furrows within one set of surface contours are defined and compared to the corresponding features in the second set of surface contours. Agreement is significant when it exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool. The statement that "sufficient agreement" exists between two toolmarks means that the **agreement is of a quantity and quality** that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility.
3. Currently the interpretation of individualization/identification is subjective in nature, founded on scientific principles and based on the examiners training and experience.

The portion appearing in bold in #2 above was inadvertently left out of: 1) the AFTE Glossary revision dated June 6, 1994, page R-1, and; 2) the reprint of the AFTE Theory of Identification published in the Winter 1998, Vol 30, No. 1 issue of the AFTE Journal, at page 86.

In accordance with the AFTE Theory of Identification, and a commitment to standardization, AFTE has developed a specific range of conclusions¹² possible when comparing toolmarks. As adopted in 1992, the range of conclusions was preceded by: "The examiner is encouraged to report the objective observations that support the findings of toolmark examinations. The examiner should be conservative when reporting the significance of these observations." These two statements were designed to give the examiner license to explain his/her reasoning for reaching his/her conclusion. These conclusions are based on a specific comparison of individual characteristics, having eliminated any possibility of subclass influence. They are:

1. **IDENTIFICATION:** Agreement of a combination of individual characteristics and all discernible class characteristics where the extent of agreement exceeds that which can occur in the comparison of toolmarks made by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.

2. **INCONCLUSIVE:**

- A.) Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
- B.) Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
- C.) Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.

3. **ELIMINATION:** Significant disagreement of discernible class characteristics and/or individual characteristics.

4. **UNSUITABLE:** Unsuitable for examination.

Scientific validity requires that the proposed theory be falsifiable by empirical testing. The AFTE Theory of Identification is plainly stated, is testable empirically using the scientific method, has been scientifically tested, has not been proven false, and, therefore, is scientifically valid.

Scientific reliability refers to the ability of the theory being tested to obtain consistent and accurate results. The AFTE Theory of Identification, based on a hypothesis that has been tested for at least 75 years^h, remains continually tested. Much of this research has been presented at professional meetings and published in peer reviewed journals. The appendix included in this paper contains only a partial list of available references. The reader is referred to *Appendix No.2 (The Application of the Scientific Method to Firearm and Toolmark Examination)* for a discussion of the use of the scientific method, in general, and in toolmark examination.

Having established that the AFTE Theory of Identification is scientifically sound, the issue of the criteria necessary to: 1) demonstrate that a tool working surface is unique and; 2) demonstrate clearly between known non-matching toolmark striae and matching toolmark striae agreement, must be addressed.

A criterion is most simply defined as a principle used as a standard in judging. Walsh and Wevers¹³ defined it this way, “As more specifically applied to toolmark identification, a criterion is some value or a threshold, which when reached or surpassed, is proof of identification of the tool as having made the mark in question.”

When a positive identification of a tool is made to the exclusion of all other tools, the examiner has observed sufficient agreement between the evidence toolmark and a test toolmark. It is the examiner’s responsibility to clearly explain what is meant by “sufficient agreement”. The AFTE Theory of Identification states that agreement is sufficient when it “exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool”.

When an examiner concludes that a tool has produced a toolmark to the exclusion of all other tools, he/she is basing this opinion on the fact that the nature of the toolmark agreement, whether impressed or striated, exceeds the best known non-matching agreement that has ever been personally observed, seen in the literature, or discussed with other examiners. In the case of impressed toolmarks, the examiner compares the agreement of the spatial relationship (location) and the contours of toolmark features transferred from a tool working surface to a questioned toolmark to individual (unique) detail in test

^h **Teale, *Popular Science Monthly*, February 1932** – The author reports on studies by Calvin Goddard of markings produced by six consecutively manufactured firing pins and four consecutively manufactured gun barrels. These studies appear to be the first documented attempt to test the firearm examiner’s ability to distinguish between consecutively manufactured tools. This is the traditional approach of “worst case scenario” testing that has been conducted by firearm and toolmark examiners to validate the ability to individualize tools. With regard to the firing pin study, Teale relates the following: “To prove his case, Col. Goddard told me he sent to the factory where the suspect’s gun was manufactured and obtained half a dozen firing pins made on the same machine one after another” and “the differences are apparent in the confirmation of the tips of six firing pins made successively on the same machine. Contact with the cartridge caps will leave identifying imprints”. With regard to the study of markings produced on consecutively manufactured gun barrels, Teale notes that “Not long ago at the Springfield Armory, in Massachusetts, bullets were fired through four rifles that had been made one after the other on the same machine. The markings on the bullets were so different that each bullet could be traced to the gun that fired it.”

marks produced by a known tool. At some point, the examiner observes a “sufficient” amount of agreement that he/she is satisfied exceeds pure coincidental agreement produced in toolmarks made from different tools, and concludes that this agreement would be seen only in toolmarks that are produced by a single tool. For impression toolmarks, this threshold is currently held in the minds eye of the examiner and is based largely on training and experience in observing the difference between known matching and known non-matching impression toolmarks. Recent theoretical research by Stone¹⁴ in establishing a numerical way to evaluate impression toolmark characteristics may very well lead to establishing evaluation criteria that can be more easily communicated from examiner to examiner and from examiner to the court. See *Appendix No .1, Section II.3 “Present Status of Impressed Toolmark Identification”* for further discussion of this issue.

All firearm and toolmark examiners require a certain amount (quantity) of agreement to establish an unequivocal identification. It is the responsibility of the firearm and toolmark examiner to know the difference between an identification and a non-identification and be able to explain the justification for their opinion. Such explanation can be difficult and is becoming more highly scrutinized. In the Ramirez case, the Florida Supreme Court reviewed testimony of five AFTE members who provided traditional justification for the identification of striated toolmarks based only on criteria established through “training and experience”. The court found that these explanations did not satisfy Florida’s rigorous Frye standards for reliability. The court therefore excluded all testimony regarding the knife mark identification and reversed Ramirez’s murder conviction. Among a number of issues was the question of the existence of some universal standard for knowing the difference between a striated toolmark identification and non-identification and, if there is such a standard, upon what scientific principles is it based. The witnesses offered no such explanation.

There is, however, a universal standard for striated toolmarks and an increasing number of toolmark examiners have taken the traditional “pattern match” process a step further by applying the conservative criteria for the identification of striated toolmarks proposed by Biasotti and Murdock.¹⁵ The method consists of counting runs of consecutive matching striae within the toolmark comparison and comparing the results to the proposed numeric threshold between an identification and a non-identification described in the conservative criteria. This conservative quantitative criteria is as follows:

1. “In three dimensional toolmarks when at least two different groups of at least three consecutive matching striae appear in the same relative position, or one group of six consecutive matching striae are in agreement in an evidence toolmark compared to a test toolmark.”
2. “In two dimensional toolmarks when at least two groups of at least five consecutive matching striae appear in the same relative position, or one group of eight consecutive matching striae are in agreement in an evidence toolmark.”

“For these criteria to apply, however, the possibility of subclass characteristicsⁱ must be ruled out.”

If the amount of consecutive agreement meets or exceeds the proposed minimum numerical criteria, the examiner is confident that the agreement represents an identification. If it does not meet or exceed the criteria, the examiner may conclude some lesser association with weighted qualifiers as described in the AFTE Range of Conclusions listed above.

Traditionally, most firearm and toolmark examiners have tabulated the quantity of agreement in their mind without consciously recording it in their notes. The AFTE Theory of Identification is based on an assessment of both quality and quantity of agreement observed between toolmarks being compared. This is how toolmark identifications have always been made. However, it currently does not require the actual recording/tabulation of the quantitative element in case notes during the examination process. There are a growing number of examiners who feel it is their responsibility to record in case notes the quantity of agreement sufficient to effect an unequivocal identification. In this manner, the agreement can be more easily communicated. Just as importantly, it is now in a form that can be compared to the work of others, providing the court with more background upon which to make an informed judgment.

ⁱ Subclass characteristics are defined as: Discernible surface features of an object which are more restrictive than *class characteristics* in that they are:

1. Produced incidental to manufacture.
2. Are significant in that they relate to a smaller group source (a subset of the class to which they belong).
3. Can arise from a source that changes over time.

Examples would include: bunter marks, extrusion marks on pipe etc. Caution should be exercised in distinguishing subclass characteristics from individual characteristics.

Based on the fact that groups/runs of consecutive matching striae constitute a means to describe a pattern, it is suggested that the quantitative element of toolmark identification was also formalized in the AFTE Theory of Identification. It is safe to say that no one has ever identified a striated toolmark without having consecutive matching striae agreement. The notion that consecutiveness is important is not new. As far back as 1932, in an article written for the *Army and Navy Journal*, Wilmer Souder, recounting ideals of (firearm identification) experts he had associated with over the previous five years, commented on the need for "... the consecutive order of the individual agreements ..."¹⁶ when describing matching bullet striations. Examiners who use the CMS tabulation approach in the interpretation of striated pattern agreement in their casework are simply recording the quantitative element of what constitutes pattern agreement in striated toolmark identification that has traditionally been kept/evaluated in the mind's eye of the examiner. Based on this characterization, it suggests that CMS is also generally accepted as a concept in toolmark identification.

With regard to the actual number separating an identification from a non-identification in a CMS protocol, the reader is invited to critically assess the studies referenced in the appendix. However, as a summary, a review of the studies involving CMS demonstrates that in testing the conservative CMS criteria for identification proposed by Biasotti and Murdock, over 4,800 actual known non-match comparisons have been conducted and tabulated and none would have resulted in a false identification. This data includes comparisons of two and three-dimensional toolmarks. Considering that these known non-match comparisons involved a wide-range of variables, and the number cited does not include in excess of a million toolmarks screened through the NIBIN system, the authors contend that sufficient support exists for the use of the conservative criteria proposed by Biasotti and Murdock. Further testing, according to the precepts of the scientific method, would certainly be welcomed.

There has been much misconception regarding the issue of consecutive matching striae (CMS), the most common being that it is somehow a "different" technique apart from pattern matching. However, it is not a "different" technique, merely an extension. When one examines a pattern, there are several elements that are part of the process of recognizing that a pattern truly exists. Whether or not the features of a striated toolmark are actually quantified or simply compared and deemed to be corresponding, several of these elements are quantifiable. For example, the location of a striation is a quantifiable entity from a fixed point such as the shoulder of a land impression. So too are the height and depth of furrows and ridges that examiners simply refer to as corresponding or "matching." So too is consecutiveness, a critical feature in defining a pattern. If it were not for consecutiveness, the striations one is attempting to compare would appear randomly and there would be no defined pattern. Indeed, no matter whether one chooses to quantify or not, the pattern remains the same. Thornton, while responding to a question raised about the validity of firearms identification, did a masterful job of describing the elements of the traditional comparison process while at the same acknowledging that "... just because the criteria for the identification of bullet striations exist largely as constructs of the mind does not mean they will remain so forever"¹⁷.

The only difference between a firearm and toolmark examiner who tabulates CMS runs in their case notes and one who does not, is the manner in which one chooses to describe the pattern. The examiner who chooses to do so without a quantifiable number is at the mercy of a language of descriptive words, words that can have different meanings to different people. In addition, it is extremely difficult, if not impossible, to describe one's concept of what constitutes a known non-match (the criterion for separating an identification from a non-identification). The authors have reviewed nearly every paper, published and non-published, with regard to identification criteria for firearms and toolmarks. Although many of these studies help to validate the science, because of the manner in which the results were communicated, descriptively, they unfortunately do little to define identification criteria.

The examiner who uses CMS does so in an attempt to better describe to others the pattern that he or she is observing. Using CMS allows for a concise and easily communicated description of a pattern that one observes. Each examiner has, or should have, an identification criterion of exceeding the best known non-match. It would appear infinitely easier to describe such a criterion in terms of the quantity of consecutive matching striae that one would expect to see in such a situation. Using such a description, two examiners could communicate something about the strength of an identification, even in the absence of photographs or the actual evidence. They could convey the strength of an identification even over the telephone.

A very key issue is that CMS has the appearance of being more objective than the traditional pattern matching. It is not being suggested, however, that this is the case because the process of observation and recognition of a pattern does not have objective elements. It does. However, CMS allows the examiner to apply a measurable criterion to the extent of pattern agreement observed in striated toolmark comparisons. Because of the manner in which the pattern is being described, it is easily defined and measured and is, therefore, testable by others using the scientific method. Therefore, one is no longer subject to only what one has seen through his or her own personal training and experience but now is able to draw on the

training and experiences of others, in addition to a plethora of studies that have put the conservative CMS criteria to the test. In a very simple sense, the reader is asked to assume the point of view of the court. Which will appear more objective? One who relies only on his or her own training and experience, or one who relies not only on his or her training and experience but draws as well on the studies and experiences of others?

More could be said regarding CMS, but it is beyond the scope of this paper. It is enough to state that CMS is not a new technique, nor in conflict with the traditional pattern matching that has characterized the discipline from the earliest of times. It is simply an extension, a manner of describing the pattern that is believed to be more concise, more easily understood, and allows for its use by others. See *Appendix No. 1, Section II.2 "CMS Theory"* for a review of cited references validating this approach.

Having established the scientific foundation for firearm and toolmark identification, an issue relevant to all legal challenges, it is time to move forward to three other relevant issues including error rate (which includes the existence and maintenance of standards and control), publication and peer review, and general acceptance.

Error Rate

Another challenge by which the court may evaluate the reliability of scientific evidence is the error rate of the method being used. This discussion addresses that challenge.

"Many considerations will bear on the inquiry, ... its known or potential error rate, and the existence and maintenance of standards controlling its operation." This is the third Daubert element that a judge may consider in determining the acceptance of scientific testimony. Frequently only the first half of this element is discussed. However, the second half should also be carefully considered. Each will receive proper attention in this section.

In our profession, we are often called to testify that a specific tool (firearms included) has been identified to a questioned toolmark. The court in these cases wants to know how often such identifications are in error. The court wants to know how often the profession, using accepted techniques and controls, produces a mistaken identity. The statement that the science of firearm and toolmark identification has a "0%" error rate is clearly not responsive to the court when questions of error rate are brought forward. The court is not interested in "theoretical error rate", which assumes everything has been done correctly and the correct answer obtained, but is interested in the real life potential error rate that is reflective of all human endeavors.

In regard to known or potential error rate within the discipline of firearm and toolmark identification, examiners should readily disclose that there have been quasi error rates reported. **To proffer that firearm and toolmark identification is "infallible" is simply not true and will be met with immediate suspicion.** The court is interested in "known or potential error rate" as a means by which to assign weight to the examiners testimony. The examiner will be more credible by readily discussing the reported error rates in the process of firearm and toolmark identification (*i.e.*, the first half of the Daubert element) and then be prepared to discuss what steps have been taken as an individual and through laboratory peer and administrative review processes to eliminate the possibility of error in the work currently being presented in court (*i.e.*, the second half of the element).

Collaborative Testing Service (CTS) is currently the only source of international proficiency testing results in the firearm and toolmark identification discipline from which a source of potential error rate may be inferred.^j Approximately 20

^j Much has been studied and reported in the literature about the fundamental principles and accuracy of comparing toolmarks to tools (firearms, etc.) known to be sequentially manufactured. However, recent studies that have a "blind" component to the comparisons, most notably comparing fired bullets from ten (10) consecutively produced barrels (Brundage, D. "The Identification of Consecutively Rifled Gun Barrels." presented at the 25th AFTE Training Seminar, Indianapolis, IN, June 1994 and Hamby, J., "Update to the Ten Barrel Test", presented at 33rd AFTE Training Seminar, San Antonio, TX, May 2002) and a knife blade study (Thompson, E. and Wyant, R. "Consecutively Made Knife Blades: Part 1". Presented at the 33rd AFTE Training Seminar, San Antonio, TX, May 2002), do show very low, near zero, error rates and it is tempting to use these error rates as being reflective of casework error rates. However, the low error rates in these studies are low because, though the examinations are laborious, the microscopic comparisons are relatively straightforward. Such comparisons are not typical of much actual casework where the toolmarks are often very limited in quantity and quality of information. These studies, therefore, do not measure the "error rate" question that the court requires. These tests do, however, support the scientific basis and accuracy of the FA/TM identification science, the fundamental theory of the science as expressed in the AFTE Theory of Identification, and the practitioners thereof.

percent of the CTS firearm and toolmark proficiency tests are sent to international subscribers. The only published document that discusses “error rate” in regard to crime laboratory proficiency testing in firearm and non-firearm toolmarks is the summary of the CTS proficiency tests by Peterson and Markham.¹⁸ Peterson and Markham looked for trends and although there is indeed some value in this information, the results are misleading and there is a strong argument that the error rate for firearm and toolmark identification casework is indeed much smaller than the error rates reported for the CTS tests.

Based on Peterson and Markham’s study of proficiency test results from 1978 through 1991, those who would tend towards an improper interpretation of the numbers could claim that a seemingly alarming 12% error rate is reported for firearms and 26% error rate is reported for toolmarks. These results, however, include “inconclusive” responses as incorrect. Grzybowski and Murdock discuss this issue in some detail when they write, “Assuming a scenario least favorable to our profession by not taking into account the number of inconclusive results (which, in fact, are neither correct or not correct), the error rate based on the total percentage of correct responses is 12% for firearms and 26% for toolmarks. But if we view the error rate strictly as a function of incorrect responses, as we believe it should be viewed since we feel that this is the error rate the judicial system is interested in, the results are far better: 1.4% for firearms and 4% for toolmarks.”¹⁹ It is important to note that the 1.4 and 4 percentages include false eliminations. If the false eliminations are not included, the percentages for false identifications are approximately 0.6% for firearms and 1.5% for toolmarks.^k

While some may balk at the use of CTS proficiency test figures, Grzybowski and Murdock suggest that, despite their limitations, “Nevertheless, the overall purpose of them is to indirectly test the validity of a particular method and protocol. We may not be altogether happy with the results, but at least a general indication of error rate is determined through these tests. It’s not the best data, but the courts are interested in this information and we should be prepared to describe the tests as well as their strengths and limitations accurately.”²⁰

As useful as the courts may find this information, it is also important to indicate that there are some limitations with respect to using this data. In their article, Peterson and Markham themselves discuss limitations of using the CTS reports citing: 1) problems of test administration which influenced test results; 2) lack of control over levels of difficulty of tests, resulting in not being able to determine if an improvement or a decline in test participation performance reflects better proficiency of the firearm and toolmark identification community or easier tests; and 3) continual changes in test design and sample make-up, making comparison of results difficult, as little replication of tests, considering these factors, was conducted to establish a baseline to monitor participant performance. The reader should be familiar with these limitations. Peterson and Markham further note that many laboratories were **obligated by policy to report inconclusive results** when the tool was not supplied. In firearms, the percentage of “inconclusives” seemed to be a function of the difficulty of the test, the clarity of instructions in the test process, and the discomfort expressed by some takers in not being supplied with the actual firearm with which they could conduct their own test firings.

The examiner should be prepared to point out other problems with using the results of CTS proficiency tests to determine error rate. One of these is that some of the test participant results in the Peterson and Markham study may not have been peer reviewed. This is because the CTS tests were supposed to represent tests of individual examiners. Also, some laboratories had different examiners share tests for internal quality control purposes. When this was done, individuals were not supposed to share their results with anyone else. On the other hand, even though ASCLD-Lab does not mandate peer review of proficiency tests, most CTS proficiency tests that have been done in ASCLD/LAB accredited laboratories since December 1997 probably are thoroughly peer reviewed because the results of select proficiency tests have to be reported to the ASCLD/LAB Proficiency Review Committee (PRC). As a result, one would think that the reported error rate for firearms and for toolmarks should be lower following PRC formation in December 1997. A recent (April 2003) study by Robert Thompson of the Peterson and Markham data and the CTS summary reports in firearm and toolmark proficiency tests for the period 1978 through 1997 (firearms), and 1981 through 1997 (toolmarks) revealed the false identification error rates to be 0.9% for firearms and 1.5% for toolmarks. The false identification error rates (post PRC) for 1998 through 2002 is 1.0% for firearms and 1.2% for toolmarks. As these are measurements only of false identifications, these percentages do not include false eliminations or “inconclusives”. Although the false identification error rates are lower, post ASCLD-Lab PRC formation, the small percentage differences are not significant. Extending this outside the realm of proficiency tests,

^k Analysis of the Peterson and Markham data by Robert Thompson, April 2003. In reviewing the CTS reports after the Peterson and Markham study, the false identification percentage was derived in the same manner as the Peterson and Markham data; the number of false identifications compared to all of the comparisons reported by the responded laboratories. For the years 1978 to 2002 the false identification rate for the firearm proficiency tests is 1.0 %, and for the years 1981 to 2002 the false identification rate for the toolmark proficiency tests is 1.3% (toolmark proficiency tests started in 1981).

Grzybowski and Murdock suggest that, “In actual casework, however, internal quality control procedures of peer and administrative review serve to further reduce the error rate.”²¹

Biasotti and Murdock²² discuss a number of deficiencies in the available literature sources that discuss forensic “error rate”. The first deficiency, and one that everyone needs to recognize, is that CTS proficiency tests were designed for laboratories to use as a quality assurance tool and not as a basis for studying error rates nationwide. They suggest that the manner in which these tests were treated may differ among participating laboratories, some choosing to treat them formally and others informally. Further, examiners that have been assigned to take these tests may tend towards more conservative results simply because there is little to gain but much to lose in the event of an error.

As others before and since, Biasotti and Murdock also discuss the limitations due to a possible lack of peer review, indicating that, “Peer review is an important process that is widely used in crime laboratories. This process helps prevent errors in casework from seeing the light of day.” Biasotti and Murdock also discuss the influence of ASCLD/LAB, suggesting that since the advent of their Proficiency Review Program in December 1997, proficiency testing was moved from a “low stakes game” to a “high stakes game.” The reason is that depending on the severity of the proficiency test result discrepancy should one be reported, sanctions could include revocation of ASCLD/LAB accreditation. Therefore, Biasotti and Murdock conclude that, “Consequently, we cannot know if pre-1998 proficiency test studies overstate or understate the accuracy of examinations.”²³ But it seems likely that post-1998 proficiency test studies may more properly reflect the accuracy of routine forensic firearm/toolmark examinations, simply because peer review will probably be done to help ensure correct results. In general, as the ASCLD-Lab accreditation process has evolved, proficiency testing has become an important part of the process and forensic laboratories are paying more attention to what they submit.

A review of the record of proficiency tests must also put these tests in the context that they are generally given; as declared tests known to the examiners as proficiency tests. Attempts to give “blind” or even “double-blind” proficiency tests in an operational crime laboratory, especially in the case of firearms/toolmark examinations, usually fail because of the laboratory’s need to communicate with their clients to obtain background information about the submitted evidence. For a more detailed practical guide through the citations designed to prepare the examiner to more fluently discuss the current status of firearm and toolmark identification error rates in court, see *Appendix No. 1, Section III “Known or Potential Error Rate”*.

The third element of Daubert also concerns peer review, which is the existence and maintenance of controlling standards of operation. The existence and maintenance of controlling standards of operation is an important concept that helps minimize analytical and interpretation errors in the crime laboratory. The use of quality control and quality assurance measures, in addition to peer review: 1) helps to maintain integrity of the work; 2) increases the likelihood of the discovery of error; 3) is necessary to minimize bias in examination and interpretation; and 4) helps bring forth a quality product. These operational standards are derived from the peer scientific community, and from forensic laboratory guidelines and policies.

So far two issues have been discussed: 1) reported summaries of firearm and incorporated into toolmark proficiency tests and; 2) what is done in most forensic laboratories to ensure that the results of individual cases are correct as reported. Next, the error rate of the individual examiner needs to be addressed. Two kinds of samples need to be considered. One type of sample is the CTS proficiency test. This error rate is easy to calculate by simply tabulating firearm and non-firearm toolmark test results. If no errors have been made, the rate is zero. If errors have been made, the error rate is calculated by dividing the number of errors of each category by the total number of proficiency samples examined. But if any errors have been made, the critical issue of why they were made must be explored. For example, the examiner may have been in a training mode and the CTS proficiency test was used merely as part of in-house training, or, the examiner may have rushed through the test, thinking that it was necessary to “get it out the door to be included in the summary CTS report”. As such, proficiency tests were never intended to be used to calculate a national error rate. Instead, they were advertised as a quality assurance tool available for use by individuals and individual forensic laboratories. CTS itself attempts to put the results of their proficiency tests in proper perspective by including the following admonition on the cover page of their summary reports:

“This report contains the data received from the participants in this test. Since these participants are located in many countries around the world, and it is their option how the samples are to be used (e.g. training exercises, known or blind proficiency testing, research and development of new techniques, etc), the results compiled in the Summary Report are not intended to be an over view of the quality of work performed in the profession and cannot be interpreted as such. The Summary Comments are included for the benefit of participants to assist with maintaining or enhancing the quality of their results. These comments are not intended to reflect the general state of the art within the profession.”

Another factor to consider is the test make up. Since large numbers of tests have to be produced, to uniform requirements, most tend to be rather straightforward and only of moderate difficulty. So, if an examiner has a zero error rate on them, is that a true measure of ability? The authors do not think that it is. Proficiency tests can be made that are so easy that everyone will get the correct answer or so difficult that very few get the correct answer. Generalizations concerning proficiency tests cannot be made unless there is some information accompanying the test that somehow quantifies the level of difficulty pertaining to the test. Since this is never done, or at least hasn't been done up to this point, proficiency test results cannot be interpreted as reflecting the universal proficiency of the profession.

Error rates associated with normal day-to-day forensic firearm/toolmark casework are impossible to establish. How is an error rate calculated, for examiners considered to be proficient, for these types of cases? The answer is, for the most part, one cannot be calculated. Since the identification sciences (firearms/toolmark, fingerprints, documents, etc.) are primarily focused on the determination of identification, that is, the proof of a singular source of evidence, a meaningful review of the historical forensic laboratory records of this type of evidence would be impossible. The scientific product is not in a form to be subjected to routine measurement, which could later be summarized, and compared to an ideal measurement. Moenssens²⁴ expressed this limitation succinctly by saying "It is very difficult to measure the probativeness of a particular examination because it deals with a unique event that cannot be statistically duplicated..." and "none of the forensic science disciplines that deal with comparisons based on unique occurrences, or clinical judgments, can establish the degree of confidence statistically with respect to an individual result." When a forensic laboratory processes a case and, following technical peer review, issues a report, the results are assumed to be correct. Most of these cases are never critically evaluated by anyone else. A small number of cases, however, are critically evaluated by qualified examiners hired by opposing counsel. If this evaluation uncovered a "legitimate error", this could, ostensibly, be used to calculate an individual examiners error rate with respect to day-to-day casework. Even in these instances, however, the sample size would be so small that a reliable error rate could not be calculated. The difficulty here is in instances where opposing experts have differing opinions about the same evidence. With subjective standards/criteria, how does one determine if either examiner has made an error? In summary, it is not possible to calculate an absolute error rate for routine casework.

So, what does this mean for the individual examiner? The examiner must first acknowledge that errors can be made. The examiner must then be prepared to discuss the CTS tests and their limitations, and recognize that, despite their limitations, they may offer the court some indication of error. It does not mean, for example, in the instance of a 1.5% CTS error rate, that every toolmark identification case report is subject to being right only 98.5% of the time, but rather that for all those respondents, 1.5% made an incorrect association. Secondly, assuming that the work has been done thoroughly and the conclusions fully supported by clear and complete notes, it is suggested that examiners advocate that it's his/her opinion that he/she has made no error in the case at hand. It is easier to convince others of this if: 1) he or she has graphically demonstrated the basis for the opinion with the use of photographs; 2) comprehensive notes have been taken that fully support the conclusions in the lab report and; 3) the examiner's work has been technically peer reviewed and administratively reviewed per ASCLD/LAB requirements (whether or not the individual's laboratory participates in this program). Such actions would serve to further minimize any reasonable chance of error in reaching a correct conclusion and will be persuasive to those in court responsible for determining the weight to be accorded the examiner's testimony.

Peer Review and Publication

The courts recognize that they are not in a position to be experts in forensic science, yet they are the "gatekeepers" of such evidence in the courtroom. In Daubert, the court stated that one of the methods that they can use to evaluate the scientific validity and reliability of a proffered technique is to determine if it has been peer reviewed by other experts in the field. Peer review is a specific process of evaluation that requires knowledge of the scientific method.

This requires the publication of the information or technique in a professional peer reviewed journal. The peer review process allows other experts within the field to evaluate: 1) the validity of the hypothesis; 2) how it was formulated and tested; 3) whether the scientific method was followed; and 4) whether proper conclusions were reached. It also encourages others to repeat the processes (replication) to further the science.

The peer review process involving publication is a formal process. The courts recognize that an informal process occurs when examiners share information, for example during seminar presentations. They also realize that publication of a technique or process does not ensure that it is accurate or generally accepted within the scientific community. The formal peer review process, however, does subject research to the light of day, and does provide a formal procedure to help curb outrageous claims or "junk science". It allows for the method/technique to be critically analyzed by the relevant scientific community. For peer review to work however, there really must be a critical and thoughtful review. Daubert/Frye/whomever wants peer review, but they expect good peer review.

Formal peer review involves two steps, a pre-publication review process and a post-publication review process. In the forensic firearm and toolmark examination community, the premier journal is the AFTE Journal, published by the Association of Firearm and Toolmark Examiners. This journal, which started out as a newsletter, has always had a peer review process. As a newsletter, the process was very informal, and was handled by the Editor and/or his designee(s). Denio²⁵, in his review of the process, covers in detail the chronology of the AFTE Newsletter/Journal peer review process. Briefly, the process evolved from being a grammar/spelling check, to contacting companies and experts in the field to verify content, to a formal process developed by the Editorial Committee. This formal process includes specific instructions for writing and submitting manuscripts, assignment of manuscripts to other experts within the scientific community for a technical review, returning of manuscripts to authors for clarification or re-write, and a final review by the Editorial Committee.

The AFTE post-publication peer review process has also been formalized to include a "Peer Review" section in the journal. This allows members or anyone else with an interest to write and comment on an article or information previously published in a journal. Once a letter is received for this section, the original author is contacted so that they may respond to the proposed comments or questions. This section should be used whenever anyone takes exception to a published article, otherwise the article resides there as an unchallenged observation in the literature.

It is noted that the AFTE Journal is not the only peer reviewed journal in which significant pertinent literature is published. Another is the *Journal of Forensic Sciences*, which is the official journal for the American Academy of Forensic Sciences, an organization that includes a number of different forensic science disciplines. This journal, and others of similar substance, utilizes a peer review process similar to that of the AFTE Journal. See *Appendix No. 1, Section V "Peer Review and Publication"* for further discussion of this issue.

General Acceptance By The Relevant Scientific Community

This paper is designed to be useful for firearm and toolmark examiners regardless of whether they face Daubert/Kumho or Frye challenges. Indeed, the authors suggest that there appears to be a blending of evidentiary standards, such that the discussions on scientific foundation, error rate/standards of control, and peer review and publication are useful in either jurisdiction.

General acceptance in the scientific community is also a meaningful addition to this blend. General acceptance means acceptance by the professional community practicing in a particular field. Arguably, forensic firearm and toolmark examination is represented by AFTE, although there are examiners who do not belong to AFTE. For legal purposes, AFTE, its peer reviewed Journal, and Glossary represent the consensus point of view of the Firearm and Toolmark community and therefore fulfills the Frye requirement for a relevant scientific community.

The Frye standard, as announced in the original 1923 decision, has undergone several changes. In many states it is referred to as Frye III or Frye IV or Frye Plus. The Frye plus has two additional requirements; they are referred to as the "prongs of Frye. These are:

1. That the procedures (laboratory practices) implementing the theory or methodology are generally accepted, and
2. That these generally accepted procedures were indeed followed accurately to yield reliable results.

Essentially, this represents the "testability" standard of Daubert. It is the responsibility of the individual examiner to demonstrate that his/her laboratory satisfied such requirements.

The history of firearms identification and court acceptance of firearms and toolmark evidence in the United States goes back over 100 years and has been the subject of numerous publications^{26 27 28 29 30 31 32 33 34 35 36}. A special contribution was made in 1931 by Kraft^{37 38} in a comprehensive two-part summary of the literature produced during the period 1919 – 1930 that dealt with (in part one) "the identification of weapons by means of the projectile and cartridge case" and (in part two) "other questions that may arise in forensic ballistics". Kraft's critical review of many articles written in German, etc. made them available to many English speaking examiners for the first time. The AFTE Theory of Identification, developed and adopted by the relevant scientific community, has provided the toolmark identification community with a theory defining and describing the approach that examiners have traditionally taken when identifying/individualizing toolmarks.

Although the history of firearm and toolmark identification is certainly relevant, it is what is happening now that is germane. Firearm and toolmark identification must continually move forward, adjust, evaluate, re-consider, etc. This is the

hallmark of a truly scientific endeavor. See *Appendix No. 1, Section IV “General Acceptance Of The Relevant Scientific Community”* for consideration of suggested courtroom responses to this challenge.

Conclusion

The discipline of firearm and toolmark identification exists within a legal system that has certain expectations. Such expectations include that it rests on a solid foundation. Included among the expectations is the recognition on the part of the practitioners that no matter the soundness of the theory, a potential for error does exist simply because this remains a human endeavor. The examiner must be prepared to admit and discuss this potential. The examiner must also be prepared to explain how this potential error is measured, and how it is managed. Further, the legal system expects reliability. Considering that the system is poorly equipped to unilaterally make this assessment, they need to rely on the practitioners themselves. Peer review, publication and general acceptance are critical elements to this process.

The individual examiner needs to discuss these issues in court in a manner that is clear and understandable. It is not a simple feat to discuss this material clearly. This paper has provided foundational discussions by which the examiner may consider four very important elements with regard to the discipline. The individual examiner needs to build upon these discussions to develop a means by which he or she can effectively communicate the essentials of these elements to the legal system.

The references and resources provided are intended to support individual examiner’s arguments and assertions. The references are only a foundation upon which to build. It is strongly encouraged that anyone seeking to produce an effective argument be fully exposed to the original works by seeking out the resources in their entirety.

Acknowledgement

The authors are indebted to: 1) John Thornton of Forensic Analytical Specialties, Inc., for doing just what he said should be expected of a peer reviewer; a critical and thoughtful review of the entire paper; 2) Charles Leete and Doug Lucas of Collaborative Testing Service, for reviewing the error rate section and offering a number of very helpful suggestions; 3) colleagues from both outside and within the Bureau of Alcohol, Tobacco, Firearms, and Explosives for useful editorial comments.

APPENDIX No. 1

Consideration of materials referenced in the following Appendix will assist the examiner in preparing for court testimony. It should be noted that this is not an exhaustive listing of references, but includes the most significant support to most efficiently increase the examiner's confidence in addressing the issues of validity and reliability of firearm and toolmark identification. It also includes suggested approaches to convincingly describe the process of examination. It is suggested that the reader assemble a binder of reference materials for periodic review and preparation for court testimony. As new references become available, the reader can keep current by updating their personal reference binder.

I. MEETING THE DAUBERT CHALLENGE - "THE ROADMAP"

1998 – Grzybowski, R., Murdock, J., "Firearm and Toolmark Identification – Meeting the Daubert Challenge." AFTE Journal 1998: 30(1): 3 –14.

This article is a dissertation describing in great detail the application of the scientific method in firearm and toolmark identification and addresses each of the four challenges of the Daubert Decision.

1997 – Biasotti, A., Murdock, J., Firearm and Toolmark Identification, Chapter 23, Vol. 2. in Modern Scientific Evidence: The Law and Science of Expert Testimony (Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds., 1st 1997 edition), St. Paul: West

This document is an excellent review of the firearm and toolmark identification process and provides, for the first time, a proposed conservative consecutive matching striae (CMS) numerical criteria for the evaluation of striated toolmark identification (that has since been validated through empirical testing). It also discusses the role of the scientific method in this discipline. It is recommended that this document be read in its entirety. Please note the cross-examination questions at the end of the chapter.

2002 – Biasotti, A., Murdock, J., Firearm and Toolmark Identification, Chapter 29, Vol. 3. in Modern Scientific Evidence: The Law and Science of Expert Testimony (Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds, 2002), St. Paul: West

This is an updated edition of the above reference containing significant references to additional validation work completed since the original publication in 1997. The conservative CMS numerical criteria remain the same as first proposed in 1997.

II. VALIDATION OF THE ABILITY TO INDIVIDUALIZE TOOLS

The following reference list is intended to provide a central resource of abstracted papers supporting the validity of a toolmark examiner's ability to individualize tools using qualitative and quantitative approaches in keeping with the AFTE Theory of Identification. The references are listed by general subject matter categories. It should be noted that the summaries reflect the personal interpretation of the authors. Those who wish to use these reference materials for their own purposes are encouraged to read the cited references in their entirety and make their own personal assessments.

The scientific method is a process of problem solving that continually contributes to a self-evolving, self-correcting body of knowledge built upon verification and validation. It is a systematic mechanism from which science develops knowledge, with principles that are clearly enunciated. Firearm and toolmark identification fits neatly into this definition. Examiners in this field practice the scientific method both formally and informally.

II.1 USE OF THE SCIENTIFIC METHOD IN VALIDATING THE ABILITY TO INDIVIDUALIZE TOOLS AS THE SOURCE OF TOOLMARKS TO THE EXCLUSION OF ALL OTHER TOOLS.

Forensic toolmark identification is a scientific discipline that is concerned with the identification of a toolmark to a specific tool. Firearms identification is a specialized area of toolmark identification dealing with firearms, a specific category of tools. The principles and theory behind this science is the same for both areas.

Firearm and toolmark examiners can establish associations of a common source by studying the difference between toolmark agreement from a common source (known matches) and comparing it to toolmark agreement from different sources (known non-matches) to establish a threshold for being able to differentiate between the two, in order to predict toolmark individualizations. Using this approach, all studies using mathematical models, mechanical models, and actual toolmarks using new and used tools have shown that a relatively small amount of matching individual characteristics can

allow us to predict the ability to individualize toolmarks to a specific source. The testing has been conducted in two ways. The first being to design “worst case scenario” models because theoretically they represent the highest probability of two different tools leaving toolmarks which cannot be distinguished from one another microscopically. An example of one model would be the comparison of consecutively produced tools such as rifled gun barrels. These models are useful as they represent the worst situations examiners might encounter in the real world.

The majority of toolmark studies were/are aimed at trying to distinguish toolmarks made by tools whose working surfaces were/are produced consecutively by various methods of manufacture. It was found, that assuming no subclass influence, this could be done. The results of these studies have made it possible to predict the ability to distinguish tools from one another based upon toolmarks made by their working surfaces. These empirical studies have been/are conducted by using the scientific method and continue to validate the theory of toolmark identification as described in the AFTE Theory of Identification. These studies have also been peer reviewed, published in the scientific literature, and are available for attempts at replication by the relevant scientific community. Select examples of this continuing research follows:

II.1.1 BULLETS AND BARRELS – EMPIRICAL STUDIES

1970 – Lutz, M., Consecutive Revolver Barrels. AFTE Newsletter #8, 1970 Aug; 24-28.

This is a study of test firings from two consecutively manufactured .38 Special revolver barrels. The author was able to identify bullets to the correct barrels and he described a lack of carry-over in land impressions but did not compare groove impressions. The method of rifling is not specified.

1975 – Skolrood, R., Comparison of Bullets Fired From Consecutively Rifled Cooney .22 Caliber Barrels. *Canadian Society of Forensic Science Journal* 1975; 8(2): 49-52.

The author did additional work on Cooney .22 caliber rifles to test the conclusions made by Churchman in May 1949, RCMP Gazette, Vol. 11, #5. No “carry-over” was noted. For Churchman’s work see AFTE Journal 1981, 13(1): 46-52.

1978 – Freeman, R., Consecutively Rifled Polygon Barrels. AFTE Journal 1978 Jun; 10(2): 40-42.

This was a three consecutively rifled H&K 9mm Luger polygonal rifled barrel study. The author noted that each barrel had its own distinct and separate individuality.

1981 – Murdock, J., A General Discussion of Gun Barrel Individuality and an Empirical Assessment of the Individuality of Consecutively Button Rifled .22 Caliber Rifle Barrels. AFTE Journal 1981 Jul; 13(3): 84-111.

This is an extensively researched paper studying gun barrel individuality and empirical assessment of button rifled barrels. The author assessed the individuality of consecutively manufactured button rifled .22 caliber rifle barrels and found each to be individual.

1983 – Hall, E., Bullet Markings From Consecutively Rifled Shilen DGA Barrels. AFTE Journal 1983 Jan; 15(1): 33-47.

Hall performed a study of four consecutively rifled barrels with polygonal rifling. The first initial test shots could not be identified to one another. (Supports rapidly changing “new barrel” phenomena). He then studied 31 test fires from each of four barrels and concluded: “with bullets closely related in the firing sequence the dissimilarity of marks created by any two different barrels is significantly greater than the dissimilarity seen on bullet pairs that are from the same barrel”.

1985 – Matty, W., A Comparison of Three Individual Barrels Produced From One Button Rifled Barrel Blank. AFTE Journal 1985 Jul; 17(3): 64-69.

This is a study of the individuality of three revolver barrels, all cut from the same section of rifled tube; “to ascertain the possibility of carry-over effects from imperfections on a rifling button”. The author examined Mikrosil casts of barrels first; then test fired. He observed no significant carry-over of subclass in casts of lands but numerous longitudinal striations on the grooves caused by button imperfections that persisted in some cases along the entire length of all three barrels. Test bullets from these barrels exhibited “sufficient carry-over for phasing in groove impressions but not enough for conclusive identification”. No subclass carry-over in land impressions was noted. Each barrel could be distinguished from one another. Matty stressed the importance of examining the barrel interior.

1994 – Brundage, D., The Identification of Consecutively Rifled Gun Barrels. AFTE Twenty-Fifth Seminar; 1994 Jun 5-10; Indianapolis, IN.

Brundage sent test fires from 10 consecutively rifled barrels to 30 labs. He also provided 15 unknowns with at least one being fired from each of the 10 barrels. All examiners successfully associated the unknowns with the barrels from which they were fired.

1995 – Brown, C., Bryant, W. Consecutively Rifled Gun Barrels Present in Most Crime Labs. AFTE Journal 1995 Jul; 27(3): 254-258.

The authors compared barrels from multi-barreled derringers and discovered that several barrels appeared to have been manufactured closely to one another with some indication of subclass carry-over on the groove impressions of bullets fired from them. No carry-over was noted on the land impression areas. These firearms represent a good source of consecutively rifled barrels.

II.1.2 CARTRIDGE CASINGS

1984 – Matty, W., Johnson, T. A Comparison of Manufacturing Marks on Smith & Wesson Firing Pins. AFTE Journal 1984 Jul; 16(3): 51-56.

The authors examined consecutively manufactured firing pins from Smith and Wesson K-frame revolvers and described the lathe-turning operation used in producing the finish. The firing pins were compared using direct comparison, lead impressions, and with impressions on fired casings. The machining process created subclass carry-over features and the authors concluded that random defects had to be present to effect an identification.

1984 – Matty, W., Raven .25 Automatic Pistol Breech Face Toolmarks. AFTE Journal 1984 Jul; 16(3): 57-60.

Matty performed an assessment of consecutively tooled Raven .25 Auto Breech faces. He described the milling process used to surface these tools and used six consecutively manufactured bolts. Three were collected and then three additional consecutively produced bolts were collected that were manufactured an hour later. His conclusion was; that each bolt could be identified to itself when offset by 180 degrees, but there was no agreement among the different breech faces. Although the tool cutting edge didn't change, micro-welded cuttings appeared to cause the changes between bolts.

1986 – Uchiyama, T., Similarity Among Breech Face Marks Fired From Guns With Close Serial Numbers. AFTE Journal 1986 Jul; 18(3): 15-52.

This is a study of breech face comparisons using firearms close in serial number (serial numbers do not necessarily correlate to consecutively manufactured firearms). The author presented excellent photographs of known non-matches depicting agreement of the coarser striated markings on breech face marks from .25 Auto Browning pistol (straight breech face toolmark finish) and .25 Auto Raven pistol (concentric circle finish). He also presents numerous known non-match comparisons that were photographically documented. Significant subclass carry-over agreement was noted among breech face toolmarks produced by these firearms.

1994 – Thompson, E., Phoenix Arms (Raven) Breech Face Toolmarks. AFTE Journal 1994 Apr; 26(2): 134-135.

This is a study of breech faces of Raven firearms that is similar to Matty's 1984 study cited above. Thompson wanted to know if there was any change in manufacturing due to change in ownership (Raven to Phoenix). He obtained four randomly selected bolts from a batch of 60,000. Each bolt had been end milled which resulted in concentric circular markings on the breech face. He compared Mikrosil casts and found, like Matty, that each bolt could be distinguished from one another.

1995 – Lardizabal, P., Cartridge Case Study of the Heckler and Koch USP. AFTE Journal 1995 Jan; 27(1): 49-51.

Lardizabal examined the correspondence of breech faces of three Heckler and Koch .40 S&W pistols. Two pistols bore sequential serial numbers and were sequentially manufactured. The breech faces were broached. Cartridge casings fired from the consecutively made pistols exhibited significant corresponding impression markings, suitable for identification, which originated from a striated subclass toolmark appearing to have been made after the breech face finishing process. This striated toolmark was located just above the firing pin hole and persisted over 250 test firings. No other marks could

be used to identify a casing to one of the two firearms. The third firearm with a lower serial number demonstrated no such striated toolmark on the breech face.

1996 – Thompson, E., False Breech Face Id's. AFTE Journal 1996 Apr; 28(2): 95-96.

Thompson cited a case example involving Lorcin 9mm Luger caliber pistols. He observed that these Lorcin breech faces were stamped and painted with heavy black paint without further finishing. He found that this results in significant family resemblance with significant possibility of misidentification based on breech face markings alone. Thompson, like Lardizabal, suggests the use of other markings, including extractor or chamber markings, to facilitate an ID.

II.1.3 TOOLMARKS

1964 – Burd, D., Kirk, P., Toolmarks: Factors Involved in Their Comparison and Use as Evidence. Journal of Police Science 1942; 32(6): 679-686.

The authors discussed individual and class characteristics of toolmarks related to screwdrivers. They intercompared three randomly selected screwdrivers of the same manufacturer and model and described what is commonly known today as subclass characteristics that they attributed to certain portions of the tool which were produced by a stamping process. Additional subclass features were noted where the edges had been sheared during the stamping process. The corners of the screwdrivers produced non-repetitive markings individual to each screwdriver. The important contribution of this work was the emphasis on the need for the examiner to evaluate the manufacturing technique used to produce tool working surfaces.

1975 – Butcher, S., Pugh, D., A Study of Marks Made by Bolt Cutters. Journal of the Forensic Science Society 1975 Apr; 15(2): 115-126.

The authors studied marks produced from bolt cutters. They described the bolt cutter manufacturing process and the comparison of known matches and known non-matches. They addressed three points related to criteria: 1) the need for assessing the quality of the toolmark (quality and contour); 2) true match vs. false match; and 3) a toolmark less than 2mm wide increases the possibility of a false match.

1975 – Reitz, J., An Unusual Toolmark Identification Case. AFTE Journal 1975 Dec; 7(3): 40-43.

Reitz studied toolmarks made from sequentially ground and non-sequentially ground twist drill bits. He concluded that toolmarks from these tools could be correctly identified to the responsible tool and that the identification was unequivocal.

1976 – Vandiver, J., New Screwdrivers Production and Identification. AFTE Journal 1976 Mar; 8(1): 29-52.

Vandiver studied the individuality of seven pairs of screwdrivers from seven different companies. The manufacturing methods were provided with each screwdriver. He concluded that marks made from these screwdrivers could be uniquely identified to the responsible tools.

1978 – Watson, D., The Identification of Toolmarks Produced from Consecutively Manufactured Knife Blades in Soft Plastics. AFTE Journal 1978 Jul; 10(3): 43-45.

Watson studied toolmarks made from two consecutively manufactured knives and observed “no carry-over”.

1980 – Cassidy, F., Examination of Toolmarks from Sequentially Manufactured Tongue and Groove Pliers. Journal of Forensic Sciences 1980 Oct; 25(4): 796-809.

Cassidy performed a detailed study and comparison of consecutively broached tongue and groove pliers. He observed no family resemblance (subclass) and he introduced the observation that the direction in which toolmarks are made would have a direct effect on the potential transfer of subclass toolmarks. An excellent description of the thought process involved in toolmark examination is presented.

1982 – Tiura, Y. J., Tire Stabbing With Consecutively Manufactured Knives. AFTE Journal. 1982 Jan; 14(1): 50-52.

This is a study of two consecutively manufactured knives. Tiura concluded that identifications were possible without any danger of confusion arising because of consecutive manufacture.

1985 – Kreiser, J., Identification of Cast Bullets and Their Molds. AFTE Journal 1985 Jul; 17(3): 88-90.

Krieser studied mold marks on lead cast bullets and considered whether or not the mold marks were unique. He acquired six different mold blocks and cast bullets from each. He observed matching subclass toolmarks in the nose areas of bullets made from different molds. This study is significant in that if one finds subclass toolmarks on undamaged evidence bullets, it would be difficult to distinguish them from truly individual toolmarks on damaged/deformed evidence bullets.

1989 – Hornsby, B., MCC Bolt Cutters. AFTE Journal 1989 Jul; 21(3): 508.

Hornsby describes the MCC bolt cutter blade manufacturing process and performed comparisons of three sets of cutting blades from one production run of 1000 blades. He concluded, “tests disclosed individual characteristics so different that there would be no possibility of misidentification”.

1991 – Warren, G., Glass Cutter Impression Identification. AFTE Journal 1991 Oct; 23(4): 925-927.

Warren studied rotary glass cutters and described the grinding operation involved in the manufacture of the cutting wheels. He concluded that the patterns produced from cutters were unique to each cutter.

1992 – Hall, J., Consecutive Cuts by Bolt Cutters and Their Effect on Identification. AFTE Journal 1992 Jul; 24(3): 260-272.

Hall studied bolt cutters and the persistence of marks made over the course of multiple cuts. He also studied toolmarks from three consecutively assembled (not manufactured) bolt cutters and concluded that toolmarks could be uniquely identified to each tool. He did not describe any observation of subclass agreement.

II.2 CONSECUTIVE MATCHING STRIAE (CMS) THEORY

There are a number of scientists in the firearm/toolmark identification community who have specifically researched the individuality of striated toolmarks based on the quantity of consecutively matching striae. These scientists approached this through the use of theoretical, mathematical and empirical studies. By doing this, the examiner is able, from data gathered using the scientific method, to assign measurable weight to the qualitative component described in the AFTE Theory of Identification. This allows the examiner to supplement his/her training and experience with data that is scientifically defensible. The following references validate this approach:

II.2.1 GENERAL

1959 – Biasotti, A., A Statistical Study of the Individual Characteristics of Fired Bullets. Journal of Forensic Sciences 1959 Jan; 4(1): 34-50.

This paper summarizes the study of the probability of occurrence of consecutive matching striae in land impressions from fired bullets in both match and non-match positions reported by the author in 1955. The findings are based on the author's research conducted between 1951 and 1955. Biasotti concluded that, “The most significant point of the data collected is the fact that 3 consecutive matching lines for lead bullets and 4 consecutive matching lines for metal-cased bullets appears to be the dividing line between data for same and different guns; and therefore, these critical series form the base line upon which the data for bullets from the same gun can be differentiated from the data for different guns.” Therefore, approximately 3 - 4 consecutive matching striae appeared to be the threshold number between a known match and known non-match.

II.2.2 MATHEMATICAL MODELS

1970 – Brackett, J., A Study of Idealized Striated Marks and Their Comparison Using Models. Journal of the Forensic Science Society 1970; 10(1): 27-56.

Brackett published his study of various mathematical models that could be applied to the study of “idealized” striated marks including geometric models, number-based models, random number outcome models, and random number replica models. His purpose was to develop a theoretical basis for striated mark analysis that could be developed into mechanical (empirical) models that could be compared to toolmarks in order to obtain sufficient information to establish objective criteria for identity between two sets of toolmarks. He discovered from these different models that the development of a random number table was quite reliable in representing a striated toolmark model. Although his ideal striations were

devoid of width, his work is significant in that he demonstrated the concept of consecutiveness alone to be a very powerful tool in deciphering coincidence from common associations. His ideal mathematical striated marks model was described as “tedious”, but it closely supported the empirical study of consecutive matching striae conducted by and reported in 1959 by Biasotti. The author suggested that such models could be more practically used with the assistance of computers in the future.

1980 – Blackwell, R., Framan, E. Automated Firearms Identification System (AFIDS): Phase I. AFTE Journal 1980 Oct; 12(4): 11-37.

Blackwell and Framan published their research into the development of automated firearms identification systems. In their concern for establishing applicable criteria for developing computerized systems that could be objectively and reliably “utilized to help establish a more factual basis for determining identity and nonidentity between two pairs of fired bullets”, the authors researched the literature. The results of this effort revealed that there was little in the literature that provided such objective criteria that could provide a “universal factual basis for establishing identity of a firearm” with the exception of the work of Biasotti reported in 1959. They observed “Biasotti has conducted research which could prove very useful to future developments in firearm identification”. To investigate this possibility, the authors conducted a simulation study of striated marks by applying Brackett’s formulas and models, which they found reliable, and in agreement with each other. Results of their work were found to be in general agreement with the results of Biasotti’s empirical studies described in 1959. They observed in the simulation study that “the results substantiated Biasotti’s hypothesis and regardless of the phase relationship of one sequence with the other, the chance occurrence of consecutive matching lines exceeding those proposed by Biasotti did not occur.”

1981 – Dienet, W., Studies of Models of Striated Marks Generated by Random Processes. Journal of Forensic Sciences 1981 Jan; 26(1): 35-50.

Dienet described his use of computer-aided studies of the degree of similarity of striated markings. Dienet addressed the problem that “a high degree of similarity between two sets of marks is not sufficient to identify a tool if it is highly probable that the similarity may occur by chance”. With this in mind, Dienet defined this problem in the form of the following question: “Given two patterns that are similar to a certain degree, what is the probability that such a similarity, or an even greater one, occurs at random?” He attempted to answer this question by constructing models based on assumptions that: “1) random processes generate the patterns on the tools; 2) different patterns on tools produced by the same machine are independent; and 3) the probability of the occurrence of a line is independent of its position”.

His paper described the calculation of probability of random occurrence of matches using actual striated toolmarks using blades of 20 shears that were ground to produce random imperfections that were then used to produce toolmarks in lead. Two sets of marks were produced with each tool and a 1.2 mm portion of each pattern was photographed and scanned into the computer. The images were prepared and evaluated with respect to line position. The “Digitized image data on 40 grinding marks were fed into a minicomputer, and the position values of the lines were determined semi-automatically.”

Idealized models were defined for an objective comparison of striated marks and then applied to the grinding mark data. Necessary conditions of the models were tested by comparing them with actual measured properties of the marks. Three different probability theory models were examined including combinatorial model, a renewal theory model, and binomial function fit model. The results of the model calculations were presented and the properties of the models were discussed. Each model presented certain strengths and weaknesses in fulfilling the above three requirements but were not entirely ideal. However, he also concluded that “numerical values computed with the aid of models permit an evaluation of the degree of similarity” and “for automation of pattern comparisons a preselection is possible, but any probability-related statements require additional studies and examinations”

1988 – Uchiyama, T., A Criterion for Land Mark Identification. AFTE Journal 1988 Jul; 20(3): 236-251.

Uchiyama sought to develop criteria for identification of land impressions using probability theory and also developed a significance level associated with this approach. This was developed because of his observations that neither the total number of matching striae or the percentage of matching striae was sufficient to establish identity. His significance level approach provided an evaluation of goodness of fit and primarily provided the probability of an accidental or random match of striae. He developed the probability equation based on actual fired bullets. Using his significance level of evaluation he observed that consecutiveness of matching striae played a principle part in indicating the identity of bullets fired from the same gun.

1992 – Uchiyama, T., The Probability of Corresponding Striae in Toolmarks. AFTE Journal 1992 Jul; 24(3): 273-290.

Uchiyama's study provides an estimate of the maximum number of consecutively corresponding lines that might be expected given the considerations of: 1) striae density; 2) critical coincidence ratio (CCR – a method of quantitatively representing how well two lines match relative to width, a range of zero to 1 with 1 being a perfect match); and 3) striae width. He developed a computer model to generate and compare striated toolmarks using these considerations. His model demonstrated that when the widths of lines were varied (as might be expected in actual striated toolmarks), the number of coincidental consecutive matching striae decreased with increasing deviations in striae width. From this experimentation he demonstrated that the additional influence of line width, in the critical evaluation of consecutive agreement, had significant additional influence on the maximum number of matching consecutive striae. When the conditions of his variables were such that the coefficient of variation for striae width was 0.9 and the CCR value was 0.8, the maximum number of coincidental consecutive matching striae was approximately 3 – 4. These results of a maximum of 3 to 4 consecutive lines that would represent a known non-match very closely agreed with Biasotti's original empirical work.

1997 – Biasotti, A., Murdock, J., in Faigman DL, Kaye DK, Saks MJ, Sanders J, editors. Modern Scientific Evidence: The Law and Science of Expert Testimony. St. Paul: West, 1997 – Chapter 23, Firearms and Toolmark Identification

This work includes, for the first time, a proposed conservative numerical criterion based on counting runs of consecutive matching striae. The purpose of this proposal was to offer a standard that could be used to distinguish between a striated toolmark identification and non-identification. This conservative quantitative criteria is as follows:

- “In three dimensional toolmarks when at least two different groups of at least three consecutive matching striae appear in the same relative position, or one group of six consecutive matching striae are in agreement in an evidence toolmark compared to a test toolmark.”
- “In two dimensional toolmarks when at least two groups of at least five consecutive matching striae appear in the same relative position, or one group of eight consecutive matching striae are in agreement in an evidence toolmark.”
- “For these criteria to apply, however, the possibility of subclass characteristics must be ruled out.”

II.2.3 EMPIRICAL STUDIES

1998 – Tulleners, F., Giusto, M., Hamiel J., Striae Reproducibility of Sectional Cuts of One Thompson Contender Barrel. AFTE J 1998:30(1): 62-81.

Tulleners and Giusto performed an extensive controlled study of the persistence of reproducible striae on bullets fired from a single Thompson Contender brand button rifled barrel with sequential removal of one inch segments to simulate comparison of striae from consecutively produced barrels. In essence, this amounted to a worst-case scenario for subclass carry-over than might be encountered by firearms examiners in the real world. These researchers used Biasotti's original objective criteria model for evaluating the agreement of the striations created on bullets test fired from each, increasingly shorter, remaining barrel. Their comparisons consisted of tabulating total number of lines, number of matching lines, and number of consecutive matching lines for each land and groove impression. This study closely correlated to work conducted by Biasotti (published in 1959) and mathematical models proposed by Brackett in 1970, providing validation of consecutive matching striae as a means of differentiating a striated toolmark identification from a non-identification. With one exception, no more than three consecutive striae were observed when bullets from increasingly shorter barrels were intercompared. “It is quite apparent that after three consecutive matching striae, one begins to move from the threshold of non-matching to a matching regime.”

The one exception was a set of six CMS that appeared to be axial and not originate from the barrel rifling. It is also significant to note that Tulleners and Giusto observed that, while there were some differences in the counts of single matching striae, the consecutive line counts by each student were very consistent, providing the first documented study of variation in counts due to “subjectivity” in striae tabulations. If this kind of consistency can be achieved with student interns utilizing strict rules for interpreting line width and contour, there are significant implications as to the potential for

consistency among trained firearms examiners if they were to follow the same guidelines for interpreting striae¹ consecutiveness.

1998 – Tulleners, F., Stoney, D. and Hamiel, J., An analysis of Consecutive Striae on Random and Consecutive Chisels. American Academy of Forensic Sciences Meeting, February 1999, Orlando, FL

Tulleners summarized a cumulative study comparing striae patterns left by six consecutively manufactured chisels and four chisels of the same size and brand, chosen randomly. The results were: “there were no consecutive runs of striae greater than 4X in known non-matching positions. They observed that “the absence of any 4X or higher occurrences in the mismatched positions, along with their routine and multiple occurrence in properly phased known matching toolmarks, provides the foundation for the hypothesis: chisel marks have an objective, measurable threshold, that can be set, above which one can be certain that a given correspondence results from a true association rather than at random.” They further concluded, “these results strongly support the hypothesis that an objective threshold, based on consecutively matching striae, can be set such that there is a clear distinction between random and matching striae correspondence.

1998 – Miller, J., Criteria for Identification of Toolmarks. AFTE Journal 1998:30(1): 15-61.

In what amounted to part one of a two part study, Miller conducted a comprehensive study that tested the falsifiability of the conservative criteria for identification using known non-matching in two and three-dimensional toolmarks generated in single land impressions on fired bullets. The degree of consecutiveness of random matching striae was recorded and modeled after the experiments conducted by Biasotti; reported in 1959. His results were then compared with results of known matches. What is significant about Miller’s work is that he employed the IBIS computer to compare approximately two million^m land impressions on bullets known to have been fired from different caliber .38 Special firearms to assist him in searching for the best known non-match candidates that he could critically evaluate for consecutiveness. The results of this voluminous number of comparisons closely correlated to Biasotti’s work in 1959, Brackett’s work in 1970, and Tulleners and Guisto’s work in 1998. When Miller compared his data to the conservative quantitative criteria for identification proposed by Biasotti and Murdock in 1997, he concluded that the “conservative quantitative criteria” “has merit, and is supported”.

2000 – Miller, J., Criteria for Identification of Toolmarks Part II Single Land Impression Comparisons. AFTE Journal 2000:32(2): 116-131.

Miller continued his work begun in 1998 by attempting to falsify the conservative numerical criteria for the identification of two and three-dimensional striated toolmarks proposed by Biasotti and Murdock by expanding his experimentation to the study of single land impressions on test fired bullets in .25 Auto, 380 Auto and 9mm Luger caliber using the IBIS data base search engine. Using IBIS, Miller intercompared a total of 34,524 land impressions in caliber .25 Auto, 92,304 land impressions in caliber .380 Auto, and 102,276 land impressions in caliber 9mm Luger among test fired bullets from different firearms. For all of his known non-match two dimensional comparisons, he observed no single group of consecutive striae exceeding 4X in caliber .25 Auto, 4X in .380 Auto and 3X in caliber 9mm Luger and no combination of groups exceeding 2X-3X-5X. He concluded that the data in Part I and II of his studies “supports the use of this criteria for an identification of striated toolmarks since it virtually excludes making false identifications”. He further concluded “caliber is not a factor in applying the Biasotti and Murdock consecutive single group criteria in establishing a threshold between an identification and non-identification for known matches and known non-matches. A strict adherence to the criteria excludes the possibility of a wrong identification, but may also eliminate some identifications where only one land impression is available for comparison”. The highest three dimensional combination group observed in one land impression, regardless of caliber, was 2x – 4x. The Biasotti and Murdock requirement for two groups of five for two dimension toolmarks, and two groups of three for three dimensional toolmarks was not seen in known non-matches, and is valid.” Finally, Miller concludes “Striations viewed in consecutive groups and combinations of consecutive groups proposed by Biasotti and Murdock is a reliable criteria that excludes erroneous identification, especially in limited striae cases, as in this study, where agreement in single land impressions only was considered.”

¹ Striae are defined as the visual image (specular reflection) generated from the tops of the striated contours (bright areas) separated by the dark shadows of the valleys (or low points) in between. The width and frequency of the striae is determined by the topography of the striated contours being examined. You count the bright “lines” separated by the dark shadows, you don’t count the dark shadows in between.

^m It is important to note that these large number comparisons were done by the automated IBIS system. It is unknown if IBIS correlations can be directly related to CMS in pattern recognition. If it can, these large numbers are significant.

Authors Note:

The number of missed identifications using the strict “conservative numerical criteria for identification” should be very low considering that false eliminations are most likely to occur in those situations where the total amount of information (striae) is very limited. In toolmark identification situations where the quality and quantity of information (striae) is great or even average, the number of consecutively matching striae will easily equal and more than likely far exceed the strict numerical CMS criteria regime required to establish a positive striated toolmark identification. Further, strict adherence to the conservative criteria should give the examiner peace of mind, knowing that in extremely marginal situations where an identification is made although the quality and quantity of information is limited, a false identification has not occurred. In these situations, the examiner is still free to qualify his/her opinion using any of the remaining choices in the AFTE range of conclusions. An inconclusive opinion, strongly worded in favor of a positive identification, can still be described in a report. It is also comforting to know that the empirical research by Miller supports the need to make further refinement within the parameters of the “conservative numerical criteria for identification”, not the opposite.

2000 – Miller, J., An Examination of Two consecutively Gang Broach Rifled Barrels and a Review of the Literature. AFTE Journal 2000:32(3): 259-270.

Miller produced striated toolmark test samples by pushing two copper jacketed bullets and two lead bullets through each barrel after Mikrosil casting each barrel. He evaluated the resulting striated marks on these bullets: 1) for the presence of subclass influence; and 2) according to the proposed conservative numerical criteria for identification of striated toolmarks proposed by Biasotti and Murdock in 1997. Miller made the following conclusions:

- “Both barrels, through an examination of the bore casts, show significant reproduction of subclass carry-over within the land impressions that are more significant than in the groove impressions”.
- “Examination of bullets pushed through barrels show that those subclass characteristics did not transfer sufficiently to bullet surface to interfere with ID of striae on bullets with correct barrel”.
- “There is a significant occurrence of consecutive groups of 2X striae in all categories with larger groups appearing in only known matches”.
- “If only single land impressions are considered, with the conservative criteria for identification applied, no erroneous identifications could be made. Some missed identifications could occur”
- “If all of available land impressions are considered when applying conservative criteria for identification, fewer missed identifications could occur, and no erroneous identifications will be made”.
- “An examination of barrels in this study shows remarkable carry-over of subclass characteristics from one barrel to the next but it is not imparted to the bullet that passes (i.e. pushed) over these surfaces”.
- “Applying the conservative numerical criteria for identification demonstrated that, after elimination of subclass influence, no erroneous identifications could be made between bullets from two consecutively rifled barrels”.

This study further validates that the use of the conservative numerical criteria for critical evaluation of striated toolmark agreement will not result in a false identification. Further, it is quite possible to effect an identification even with the presence of subclass carry-over on the tool working surface(s) (in this case rifled gun barrels) as long as the significance of this influence has been sufficiently evaluated.

2001 – Miller, J., An Examination of the Application of the Consecutive Criteria for Identification of Striated Toolmarks Using Bullets Fired from Ten Consecutively Rifled Barrels. AFTE Journal 2001:33(2): 125-132.

Miller intercompared bullets test fired from ten consecutively manufactured gun barrels to determine if: “1) if they could be identified to the barrel which fired them; 2) if the presence of similar toolmarks on bullets fired from different barrels could be attributed to subclass characteristics; and 3) the propriety of the proposed conservative criteria for identification.” Miller concluded, “although some striae present in the land and groove impressions of the bullets fired from the consecutively rifled barrels could be the result of subclass influence, none of these features affected the correct identification of the bullets. None of the areas examined between different bullets were of sufficient quality to lead to a misidentification. In fact, it was difficult to find areas that could be considered as having been produced by a subclass source.” By “applying the Biasotti and Murdock conservative numerical criteria for toolmark identification, a significant difference can be seen between a known match and known non-match comparison. Considering the results of the data for the two and three-dimensional comparisons between known matches and non-matches, no erroneous identifications would be expected, although some actual identifications would be excluded.” This study further validates that the use of the conservative numerical criteria for critical evaluation of striated toolmark agreement will not result in a false identification.

II.2.4 REPORTED USE OF THE CMS CONSERVATIVE NUMERICAL CRITERIA IN CASEWORK

The following references illustrate the practical use of CMS:

2000 – Moran, B., The Application of Numerical Criteria for Identification in Casework Involving Magazine Marks and Land Impressions. AFTE Journal, 2000:32(4): 326 - 331. (reprinted in AFTE Journal, 2001:33(1): 41 – 46 due to editor's printing error in original publication)

The author describes the practical application of the conservative numerical criteria for identification in casework involving magazine marks and rifling impressions. He also describes his line counting method in step-by-step detail including use of this criteria in evaluating magazine marks on fired cartridge cases and its utility in critically evaluating those markings with very limited information (at higher risk for misinterpreting the influence of subclass features). He also discusses the importance of evaluating the tool working surface for subclass influence using magazine marks as an example.

2001 – Doelling, B., Comparison of 4000 Consecutively Fired, Steel-Jacketed Bullets. AAFS Annual Meeting, February 19-24, 2001, Seattle, WA.

In this abstract, the author briefly describes his research of the intercomparison of striated toolmarks on 4,000 consecutively fired bullets from the same barrel. He reports his findings as follows; "After 4000 rounds, it is still possible to identify a projectile back to the baseline as per Biasotti and Murdock's recommendation of a minimum of one six line run or two three line runs for identification". The real significance of this work is that any interested reader of this abstract will know exactly the criteria that Doelling used to evaluate the data reported in his experiment. This threshold offers a great advantage in communicating the basis for a striated toolmark identification by using a universal language to describe observations.

2003 – Moran, B., Comments and Clarification of Responses From A Member of the AFTE 2001 Criteria For Identification of Toolmarks Discussion Panel. AFTE Journal, 2003:35(1): 55 – 65

This paper recaps opening comments presented at the AFTE 2001 Criteria For Identification of Toolmarks Discussion Panel and: 1) expands upon the utility of CMS referred to in Moran's opening remarks; 2) qualifies Moran's statements that objections to the use of CMS can be overcome; and 3) follows up Moran's responses and concerns expressed by other participants that may not have been made clear during the discussion regarding his approaches to casework using CMS. It is suggested that the interested examiner review this paper to be better prepared to respond to such questions.

II.3 PRESENT STATUS OF IMPRESSED TOOLMARK IDENTIFICATION

If the examiner is testifying to an impressed toolmark, he/she should readily admit that currently there are no benchmark quantitative criteria like those that have been developed for striated toolmarks, although progress is being made in this area. The examiner will offer their opinion of impressed toolmark identifications using the traditional approach, admit that their opinion is based on their training and experience, explain the scientific reasoning behind impressed toolmark identification theory (agreement that exceeds the best known non-matching impression toolmarks must be present), demonstrate the identification visually with photographs (supported by thorough documentation in their notes) and impress upon the court that their work is both peer and administratively reviewed in an attempt to eliminate error. This is the best that can be offered in impressed toolmark identification and should be defended rigorously. The reader is encouraged to review the following reference in regard to the treatment of impressed toolmarks.

2003 – Stone, R., How "Unique" Are Impressed Toolmarks?, submitted for publication in the AFTE Journal, February, 2003.

This paper: 1) describes theoretical classes of impressed toolmark characteristics; 2) proposes mathematical models for these characteristics; 3) provides probability estimates using the proposed model to quantify them; and 4) based upon conservative assumptions, the probability of occurrence of combinations of such characteristics are presented. Although this paper is theoretical, it has great value in that it provides the firearm and toolmark identification community with definitions for "standardized individual characteristics" that could act as a basis for a universal language to be used by the community. Further, the paper provides guidance for others to design research experiments to begin empirically testing quantitative thresholds for deciphering between random impressed toolmark agreement and agreement that can be used as the basis for a true association to the tool that produced it, like that now enjoyed for striated toolmarks.

III. KNOWN OR POTENTIAL ERROR RATE

The following documents should be reviewed in detail to support responses to this issue:

1995 – Peterson, J., and Markham, P., Crime laboratory testing results, 1978 – 1991, II: Resolving Questions of Common Origin. Journal of Forensic Sciences 1995: 40(6): 1009 – 1029.

This is the only document that discusses “error rate” in regard to crime laboratory proficiency testing involving firearm and toolmark identification. Included in this lengthy paper are studies of toolmark identification and firearms identification. The paper summarizes, in a general way, proficiency tests conducted between 1978 and 1991.

The reader is referred to pages 1018 – 1019 for results of firearms proficiency testing and pages 1023 – 1025 for toolmark proficiency results.

1998 – Murdock, J, and Grzybowski, R., Firearm and Toolmark Identification – Meeting the Daubert Challenge. AFTE Journal 1998: 30(1): 3 –14

Refer to page 9 of this document and review the section entitled “Criterion No. 2 – Known or Potential Error Rate” and note the authors response to the results of the testing described in the above report.

1997 – Biasotti, A. and Murdock, J. Firearm and Toolmark Identification, Chapter 23, Vol. 2. in Modern Scientific Evidence: The Law and Science of Expert Testimony, Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds., (2nd ed, 2002 Vol. 3), St. Paul: West

Refer to pages 143 – 144 (1997 ed) and/or 508 – 510 (2002 ed) and read the section entitled “Disagreement Among Practitioners in Particular Applications”. A well-developed discussion of error rate is presented with a number of good references for further exploration by the interested reader. Additional information about the ASCLD/LAB Proficiency Review Program is included in the 2002 edition.

2002 – Thornton, J., Peterson, J., The General Assumption and Rationale of Forensic Identification, Modern Scientific Evidence: the Law and Science of Expert Testimony, Chapter 24, Vol. 3., p 192 –194 St. Paul: West, 2nd ed 2002

The authors provide a discussion of the relative degree of subjectivity of firearm and toolmark identification and it’s vulnerability to attack in light of Daubert criteria. Numerous references to the CTS testing results are made. This information should prove very useful to anyone preparing for a Daubert or Frye hearing.

IV. GENERAL ACCEPTANCE BY THE RELEVANT SCIENTIFIC COMMUNITY

Toolmark identification has been used in one form or another for the last 100 years as is evidenced by the numerous publications cited in this paper. The AFTE Theory of Identification has provided the toolmark identification community with a formal theory defining and describing the approach that examiners have traditionally taken when identifying/individualizing toolmarks. The quantitative element of toolmark identification and the qualitative element were both formalized in this theory. Examiners who use the CMS tabulation approach to the interpretation of striated pattern agreement in their casework are simply recording/tabulating the quantitative element of what constitutes pattern agreement in striated toolmark identification that has traditionally been kept in the mind’s eye of the examiner. If the examiner is asked if an identification using quantitative CMS is an accepted procedure in the firearm and toolmark identification community, he/she can respond as follows:

All firearm and toolmark examiners require a certain amount (quantity) of agreement to establish an unequivocal identification, and consecutiveness is a major factor in striated toolmark identifications that can be quantitated. It is the responsibility of the firearm and toolmark examiner to know the difference between an identification and a non-identification as stated in the AFTE Theory of Identification, adopted in 1992. Traditionally, firearm and toolmark examiners have tabulated the quantity in their minds eye without consciously recording it in their notes. There are a growing number of examiners who record the quantity of agreement in their case notes. They do this because they feel it makes it easier to demonstrate the reliability of striated toolmark identifications to any interested party. This process can be compared to counting rungs on a ladder. As you climb higher on the ladder, you move towards an identification. There is some point, as you climb, that you cross between a non-

identification and an identification. Toolmark examiners must know the number of rungs they need to climb to reach this point.

For those examiners who do not record runs of CMS in striated toolmark examinations, and are asked if this approach is an accepted procedure in the firearm and toolmark identification community, he/she may respond as follows:

The AFTE Theory of Identification is based on an assessment of both quality and quantity of agreement observed between toolmarks being compared. However, although CMS is a major factor to be considered in any striated toolmark comparison, the AFTE Theory of Identification does not require the actual tabulation and recording of CMS runs during the examination process. During the examination, I observed both quality and quantity of agreement that in my opinion exceeds the best known non-match that I have ever personally observed, seen in the literature, or discussed with other examiners. Therefore, based on this knowledge, in combination with my training and experience, the amount of agreement I observed in this case is sufficient for me to identify the toolmarks I examined.

This explanation, however, may not be as persuasive to the court as the tabulation of CMS runs. In addition, if an examiner does not have a keen appreciation for the extent of striae agreement that can be seen in known non-matches, misidentifications can occur when matching CMS agreement is not counted and tabulated. Indeed, in cases where limited striae are available for comparison, there is the greatest potential for differences of opinion and misidentifications, based on variations in each examiners personal criteria for identification.

V. PEER REVIEW AND PUBLICATION

It is important for the examiner in court to explain that the published data defining/supporting the process of firearm and toolmark examination has been/is peer reviewed. The AFTE Journal meets the requirement of a peer-reviewed publication. The AFTE Journal is a scientific journal published quarterly that has a formal pre-publication evaluation process, an editorial committee with subject matter experts who function as referees, and a post-publication peer review process designed to ensure that relevant and reliable information is provided to the criminal justice community and members of the discipline of Forensic Firearm and Toolmark Identification. It can be pointed out that the AFTE Journal has been historically peer reviewed by various means. The examiner is referred to the following paper to provide documentation of the history of the AFTE Journal peer review procedures:

2002- Denio, D., Forensic Firearm and Toolmark Identification, Daubert and Peer Review. AFTE Journal, 2002: 34(2): 210 – 214

This document chronicles the history of the AFTE Journal peer review process. A detailed account of the peer review procedures is described relative to specific time frames. The value of this paper is that the examiner can correlate the peer review process with the date of publication of any paper(s) he/she has relied on in offering their opinion.

VI. DESCRIBING THE PROCESS OF FIREARM AND TOOLMARK IDENTIFICATION

Examiners are often asked to describe the progression of steps in the toolmark identification process. Since these steps are practiced on a daily basis, to the point where it becomes second nature, this might seem like a simple task. However, explaining these steps to a jury can become awkward, especially if the examiner has not given thought to providing a simple but accurate description of the process. The following is a suggested outline of the essential steps in any toolmark identification process and can be used as a guide in developing responses to this type of question.

The Steps In The Process

1. Initial examination of evidence and evaluation of questioned toolmarks for:
 - a. Trace evidence
 - b. Type and configuration of tool (or firearm) used— class characteristics
 - c. Potential value for comparison and identification
2. Examination for trace evidence on tools (or firearms) submitted for comparison
3. Evaluation of class characteristics of tools (or firearms) to determine if they agree with questioned toolmark(s)
4. If class characteristics agree, test toolmarks are made
 - a. Easy with firearms since ammunition is cycled in predictable manner by the firearm
 - b. Non-firearm toolmarks may require a series of test marks produced at various angles and pressures in order to duplicate the appearance of question toolmarks

5. Prepared test marks, having the general appearance of the questioned toolmark(s), are intercompared with each other to determine if the tool is able to leave reproducible marks. If it does, the examiner can expect to be able to identify the tool as having produced the questioned toolmark(s) as long as:
 - a. The tool is indeed the source of the questioned toolmark
 - b. There has not been significant change to the tool working surface(s) since the questioned toolmark was produced
 - c. The questioned toolmark is identifiable (has sufficient, unique impression or striated markings that could potentially be positively identified with a single tool)
 - d. The working surface(s) of the tool are individual (not merely class or subclass characteristics)
6. Evaluation of the working surfaces of the tool for subclass (family) influence
 - a. If absent, then individual
 - b. If present
 - i. Are they transferred to the test marks?
 1. If no, there is no influence
 2. If yes, is there sufficient influence to preclude individualization of the tool?
 - a. If no, tool can be identified as the source of the mark to the exclusion of all others despite the presence of subclass characteristics among individualizing characteristics
 - b. If yes, tool can only be identified as a possible source within a limited group of tools that share the subclass characteristics
7. Compare test marks with the questioned toolmark
8. Make final conclusion(s) based on the above observations

VII. APPROACHES TO PROVIDING EXPERT TESTIMONY IN COURT

Numerous references list questions that the examiner should be familiar with, but few provide suggested responses. The following references are exceptions and can be referred to by the interested reader to provide ideas as to how to most effectively explain and/or defend the steps in the toolmark identification process.

Biasotti, A., Murdock, J., “Firearms and Toolmark Identification”, Appendix II, Chapter 23, Section 23-2.0, Vol 2, pp 151-153 in Modern Scientific Evidence: The Law and Science of Expert Testimony (Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds., 1997), St. Paul: West (2002 Edition, Vol. 3, pg 521-523)

The appendix section of this reference offers some excellent questions (with responses) designed to test the witness’s ability to identify toolmarks. This publication is specifically written for the legal community and therefore is readily available for any attorney wishing to closely examine the toolmark identification expert. Examiners are well advised to be familiar with the material in this appendix.

Moran, B., Firearms Examiner Expert Witness Testimony: the Forensic Firearms Identification Process including Criteria for Identification and Distance Determination, AFTE, Vol. 32, No 3, Summer 2000

This paper provides mock testimony describing the process of firearms identification. It includes an explanation of the basis for toolmark identification including the adoption of the conservative numerical criteria for striated toolmarks first proposed in Chapter 23 of Modern Scientific Evidence (1997) described in the first section of this paper. The testimony provides a logical development of the firearm identification process of both bullet and cartridge case comparisons [which can be equally applied to general (non-firearm) toolmark identification principles as well]. Included in the testimony is a discussion of the importance of evaluating tool working surfaces for subclass influence. Interested readers can obtain ideas about how to best formulate their own explanations of the identification process.

APPENDIX No. 2

THE APPLICATION OF THE SCIENTIFIC METHOD TO FIREARM AND TOOLMARK EXAMINATION

Zen And The Art of Motorcycle Maintenance – Contribution To Forensic Science – An Explanation of The Scientific Method

Bruce Moran and John Murdockⁿ

The firearm and toolmark examiner may be asked in court whether he or she applies the scientific method in the course of their routine casework. It may also be suggested that conducting such work cannot be scientific if the examiner does not have a science degree. The examiner must be able to respond to this type of questioning.

These questions may seem disconcerting to the examiner who does not have a formal science background (they may even be disconcerting to most persons having a science background). Can non-scientists conduct work in a scientific manner? The answer is YES! Firearm and toolmark examiners, whether they have a science degree or not, apply the scientific method in the pursuit of gaining knowledge when they solve problems in the every day course of their work. Even though some firearm and toolmark examiners are not formally trained as scientists, Daubert and rule 702 do not require experts to be scientists. However, it is extremely important in light of Daubert that he or she: 1) fully understand that the discipline is founded on the scientific method; and 2) has a full understanding of this concept both in theory and in practice. When non-scientists realize the simplicity of the scientific method, confidence will be gained in explaining and defending the process of firearm and toolmark identification as a science. It is hoped that the following will provide the reader with a better understanding of the scientific process used when casework is conducted and provide a suggested approach to a simple explanation in the courtroom.

Traditionally, the evaluation process involved in forensic firearm and toolmark identification is based on individual expertise built upon training and experience, and is founded on formally developed scientific principles developed from basic research. Members of this profession, while doing routine casework, are engaged in applied science. Examiners are generally not concerned with the formal basis for how the scientific method was used to develop identification theory. However, in the everyday course of their duties, they address such questions/problems as:

1. Did this gun fire this bullet?
2. How far was the muzzle of the gun from the victim when the fatal shot was fired?
3. Where was the shooter located when the fatal shot was fired?

Finding solutions to these types of routine problems does not require classical “basic research” using the scientific method. This might lead one to think that we are not applying the scientific method when conducting this type of routine casework. Indeed, we are applying the scientific method but, for the most part, not in a formally structured way.

An excellent discussion of the use of the scientific method is found in an unlikely book entitled Zen and the Art of Motorcycle Maintenance by Robert M. Pirsig^o. Pirsig’s book is an introspective look at life while traveling on a motorcycle across the USA. During the trip, Pirsig shares with the reader his philosophy on life. In chapter nine, Pirsig does an excellent job of discussing inductive and deductive reasoning and the scientific method. During the process of figuring out the solution to an electrical problem on his motorcycle, he uses the scientific method. It is an especially effective chapter because it illustrates how un-complicated the scientific method really is. We use it, informally, in our everyday life. While doing so, we don’t think consciously about the steps and we don’t keep lab notes detailing the steps in the process.

While reading chapter nine, the scientist and non-scientist alike will easily understand the concepts, and remain entertained while doing so, because of Pirsig’s writing style and use of humor. The reader will soon realize that he/she uses the scientific method almost every day but just doesn’t realize it. Pirsig also points out that while most day to day problems do not require the application of the full blown scientific method, it is there to use when really tough problems are encountered. How many of us have unconsciously used the scientific method to fix a lamp that won’t light when you turn the switch? Using Pirsig’s model, the authors offer examples of how to use the steps in the scientific method, to: 1) repair a

ⁿ B. Moran, Sacramento County District Attorney, Laboratory of Forensic Services, Sacramento, CA, (916) 874-1757
J. Murdock, Bureau of Alcohol, Tobacco, Firearms and Explosives Forensic Science Laboratory – Walnut Creek, CA – (925) 280-3624

^o Pirsig, R. M., Zen and the Art of Motorcycle Maintenance, Bantam Books, 1974, pp. 92-97.

lamp; 2) answer a routine casework problem in firearm and toolmark cases; and 3,4) solve basic research issues, thereby validating our ability to individualize toolmarks.

SCIENTIFIC METHOD	LAMP REPAIR	APPLIED SCIENCE (Routine Case Work)	BASIC RESEARCH IN: 1) Distinguishing tools from one another based on working surface produced toolmarks -(allows for prediction) and; 2) comparing the extent of agreement in 1) with toolmarks made by the same tool working surface - (allows for prediction)	BASIC RESEARCH IN: Toolmark Identification Criteria Using Tabulated Runs of CMS (Allows For Prediction)
Define a problem.	Why won't the lamp turn on when the switch is turned?	Did this gun fire this bullet?	Can tools be distinguished from one another based on toolmarks produced by their working surfaces, and will there be a qualitative/quantitative difference between these toolmarks and toolmarks produced by the same tool working surface?	When I have a small amount of matching striae in a 3D toolmark that has been produced with no subclass influence, how do I know whether I have a positive identification?
Formulate a hypothesis / tentative explanation.	The bulb has burned out.	No it did not.	Assuming no subclass influence, both randomly and consecutively manufactured tools can be distinguished from one another based upon toolmarks produced by their working surfaces, and, there will be a qualitative/quantitative difference in the extent of agreement between these toolmarks and a series of toolmarks produced by the same tool working surface.	The presence of three consecutive matching striae (3X) in a 3D toolmark, that has been produced with no subclass influence, is sufficient to establish a positive identification.
Perform experiment to test the hypothesis.	Replace it. with a new bulb.	Identify the caliber of the firearm and the questioned bullet and compare the two.	Microscopically compare toolmarks produced by the working surfaces of both randomly and consecutively manufactured tools and note the extent of the qualitative/quantitative agreement. Compare this agreement with the agreement found when comparing toolmarks made by the same tool working surface.	Compare known non-matching 3D toolmarks and record the maximum amount of consecutive (when striae are adjacent to one another) matching striae.

(TABLE continued)

SCIENTIFIC METHOD	LAMP REPAIR	APPLIED SCIENCE	BASIC RESEARCH	BASIC RESEARCH
Predicted outcome of the experiment.	If the lamp will not turn on because the bulb is burned out, it will turn on when the bulb is replaced.	If the gun is a different caliber than the bullet, the gun did not fire it.	Randomly and consecutively manufactured tools can be distinguished from one another based on their toolmarks. Toolmarks produced by the same tool can be identified with one another.	If the maximum number of consecutive matching striae in a known non-matching toolmark is two (2X), a toolmark having 3X consecutive matching striae can be identified.
Observed results of the experiment.	The new bulb doesn't turn on.	The gun is chambered for 9mm Luger and the bullet is also 9mm Luger caliber.	Randomly and consecutively manufactured tools produce toolmarks that exhibit <i>limited</i> microscopic agreement. The best such matching agreement is <i>less</i> than the microscopic agreement between toolmarks produced by the same tool.	The maximum number of consecutive known non-matching toolmark striae in 3D toolmarks is four (4X).
Conclusions from the results of the experiment.	A burned out bulb is not the cause of the lamp failure.	Since the gun and the bullet are of the same caliber, the gun cannot be excluded as having fired the bullet. The hypothesis is therefore proven false.	1) Assuming no subclass influence, the working surfaces of all tools produce discernibly different toolmarks even though <i>some</i> quality/quantity of microscopic agreement may be present. These toolmarks are referred to as known non-matches. 2) Toolmarks produced by the same tool working surface (referred to as known matches) can be identified with one another and exhibit a <i>greater</i> quality/quantity of microscopic agreement than known non-matching toolmarks. The hypothesis is supported by the experimental results.	3X consecutive matching striae agreement is less than the maximum consecutive matching striae agreement observed in the best known non-match. The hypothesis is therefore proven false.
Form a new hypothesis / tentative explanation.	The lamp will not turn on because it is not plugged in.	If the gun and bullet are the same caliber, the gun did not fire the bullet because the rifling class characteristics are different.		The presence of five consecutive matching striae (5X) in a 3D toolmark that has been produced with no subclass influence is sufficient to establish a positive identification.

(TABLE continued)

SCIENTIFIC METHOD	LAMP REPAIR	APPLIED SCIENCE	BASIC RESEARCH	BASIC RESEARCH
Perform experiment to test the hypothesis.	Check to see if the lamp is plugged in; plug the lamp in if it is not.	Compare the rifling impression class characteristics on the test bullets to the questioned bullet.		Continue comparing known non-matching 3D toolmarks and record the maximum amount of consecutive matching striae agreement.
Predicted outcome of the experiment.	If the lamp will not work because it is not plugged in, it will work when it is plugged in, assuming there is electric current.	If the rifling class characteristics are different, the gun did not fire the bullet.		If the maximum number of consecutive matching striae in a known non-matching 3D toolmark is four (4X), a toolmark having 5X consecutive matching striae can be identified.
Observed results of the experiment.	The lamp works when it is plugged in.	The test bullet exhibits five land and groove impressions and the questioned bullet has six.		The maximum number of consecutive known non-matching toolmark striae in a 3D toolmark is still four (4X).
Conclusions from the results of the experiment	The lamp was not working because the lamp was not plugged in. The new hypothesis is supported by the results of the experiment.	The gun did not fire the questioned bullet. The new hypothesis is supported by the results of the experiment.		5X consecutive matching striae agreement is greater than the maximum consecutive matching striae agreement observed in 3D known non-matching toolmarks, therefore, 5X consecutive matching striae is sufficient to establish a positive identification in 3D striated toolmarks. The new hypothesis is supported by the experiment results.

These examples include only one or two cycles of the scientific method process. By adding more boxes to the table, the process can continue as long as is needed to resolve problems, questions and inquiries that we commonly encounter in the discipline of firearm and toolmark identification. The study of this table, and its possible use in the courtroom as a demonstrative illustration of how we conduct our work, may go a long way to providing a convincing response to the

challenges posed at the beginning of this discussion. Thornton^p distinguishes between the scientific method and science by saying that the former generally leads to the latter. He accurately characterizes the forensic science endeavor by saying:

“Little of what goes on in forensic science resembles the classical description of how science develops theories, tests hypotheses, and revises its ideas and understandings. That is partly because the scientific method is a description of pure, or basic, science (knowledge building), while forensic science is an *applied* science. Ideally, forensic scientists would apply the best knowledge borrowed from basic research that *has* employed the scientific method”

One of the characteristics of a science is that it consists of an orderly body of knowledge with principles that are clearly enunciated. The authors feel that the basic research described, in a general way, in the first basic research column, and summarized by Nichols’^q, constitutes a firm scientific basis for being able to distinguish between toolmarks made by different tools and identify toolmarks made by the same tool, to the exclusion of all other tools. The authors also believe that the basic research in tabulated runs of CMS, in combination with qualitative agreement, also constitutes a firm scientific basis for numerical criteria for the identification of striated toolmarks. Nichols^{rs} has also summarized this research.

The first “basic research” column is really a summary of numerous separate studies conducted over many years. Each of these studies could be described in a “basic research” column of its own. As mentioned, Nichols has summarized the results of many of these studies. The AFTE Theory of Identification^t stemmed directly from this body of basic research.

The reader is cautioned that the second “basic research” column in the above table is used only for illustrative purposes to explain the use of the scientific method to pursue one issue in the development of a quantitative criteria for deciphering between toolmark identifications and non-identifications. The conclusion stated at the end of this illustration that “Observing agreement of 5X is sufficient to establish a positive identification in 3D striated toolmarks” is misleading as it only illustrates one consideration in the development of the validation of the conservative quantitative criteria, described below, that is presently used by many in the firearm and toolmark identification community. For example, the scientific method was used to test hypotheses addressing the issue of a quantitative threshold that could be used to identify 2D striated toolmarks. Using the same steps illustrated in the above table, research showed that observing agreement of 6X is insufficient to establish a positive identification in 2D striated toolmarks.

The scientific method was then applied to test the validity of the “conservative criteria” proposed by Biasotti and Murdock in 1997^u. The method consists of counting runs of consecutive matching striae and comparing the results to the numeric threshold between identification and non-identification proposed in the conservative criteria. This conservative quantitative criteria is as follows:

1. “In three dimensional toolmarks when at least two different groups of at least three consecutive matching striae appear in the same relative position, or one group of six consecutive matching striae are in agreement in an evidence toolmark compared to a test toolmark.”
2. “In two dimensional toolmarks when at least two groups of at least five consecutive matching striae appear in the same relative position, or one group of eight consecutive matching striae are in agreement in an evidence toolmark.”

^p Thornton, J., “The General Assumptions and Rationale of Forensic Identifications”, Chapter 20 Part IV - 1997, Chapter 24 Part IV (2nd Edition – 2002 – co-authored with Peterson, J.), Modern Scientific Evidence: The Law and Science of Expert Testimony

^q Nichols, R. Firearm and Toolmark Identification Criteria: A Review of the Literature, Journal of Forensic Sciences 1997 May; 42(3): 466-474.

^r Nichols, R. Firearm and Toolmark Identification Criteria: A Review of the Literature, Part II, Journal of Forensic Sciences. 2003 March; 48(2):318-327.

^s Nichols, R., Firearm and Toolmark Identification: The Scientific Reliability and Validity of the AFTE Theory of Identification, draft 4-13-03, the author plans to submit this to the AFTE Journal for publication.

^t The Theory of Identification, Range of Striae Comparison Reports and Modified Glossary Definitions – An AFTE Criteria for Identification Committee Report, AFTE, Vol. 24, No. 3, July 1992, pp 336 - 340

^u Biasotti, A. and Murdock, J., “Firearms and Toolmark Identification”, Chapter 23, Section 23-2.0, Vol. 2, pp 131-151 in Modern Scientific Evidence: The Law and Science of Expert Testimony (Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds., 1997), St. Paul: West, (Chapter 29 in 2nd Edition 2002)

“For these criteria to apply, however, the possibility of subclass characteristics must be ruled out.”

This proposed numerical criteria is conservative in that the quantity of consecutive striae runs required is greater than the threshold of four (4X) consecutive matching striae in three dimensional striated toolmarks described by Biasotti^v in 1959, five (5X) consecutive matching striae in two dimensional toolmarks as reported by Murdock^w in 1985 and six (6x) consecutive matching striae in two dimensional toolmarks reported by Miller^x in 1998. The proposed conservative threshold was based on a combination of six years of cumulative California Criminalistics Institute Toolmark Criteria For Identification class exercises completed by students, personal criteria developed and pooled together, and casework and studies conducted between 1957 and 1997, and was deliberately set greater than the greatest 2D and 3D known non-matching agreement to be conservative.

These conservative guidelines have been supported by much empirical testing since 1959 by studies using the scientific method process illustrated in the table above^{y z aa bb cc dd}. It has been found that to date, when using this criteria, no consecutive runs exceeding the proposed thresholds have been observed and therefore, no false identifications will be made using it. There is a possibility, however, when dealing with a very limited quantity of striae, that a true identification may be missed.

Although there is sufficient validation testing allowing toolmark examiners to use this conservative consecutive matching striae threshold, in the true spirit of the scientific method, as empirical testing continues to try and disprove this identification theory, there may be some small shift in the “known non-match” threshold observed. Indeed, it may be possible, with continued testing and increased data, to “tighten up” the proposed numerical threshold to eliminate the occasional “missed” identification.

^v Biasotti, A., A Statistical Study of the Individual Characteristics of Fired Bullets, JFS, Vol. 9, No 1, Jan. 1959

^w Biasotti, A. and Murdock, J., Criteria for Identification – State of the Art of Firearms and Toolmark Identification, AFTE Journal, Vol. 16, No 3, July 1985

^x Miller, J. and McLean, M., Criteria for Identification of Toolmarks, AFTE Journal, Vol. 30, No 1, 1998, pp. 15 - 61

^y Ibid

^z Tulleners, F., Giusto, M., and Hamiel, J., Striae Reproducibility of Sectional Cuts of One Thompson Contender Barrel, AFTE Journal, Vol. 30, No. 1, 1998, pp 62-81

^{aa} Tulleners, F., Stoney, D. Hamiel, J., An Analysis of Consecutive Striae on Random and Consecutive Chisels, Abstract - American Academy of Forensic Sciences Meeting, Orlando, FL, February 1999

^{bb} Miller, J., Criteria for Identification of Toolmarks – Part II, AFTE Journal, Vol. 32, No. 2, 2000, pp. 116 - 131

^{cc} Miller, J., An Examination of Two Consecutively Rifled Barrels and a Review of the Literature, AFTE Journal, Vol. 32, No. 3, 2000, pp. 259 - 270

^{dd} Miller, J., An Examination of the Application of the Conservative Criteria for Identification of Striated Toolmarks Using Bullets From Ten Consecutively Rifle Barrels, AFTE Journal, Vol. 32, No. 2, 2001, pp 125 - 132

References

¹United States v. Plaza, Nos. CR. 98-362-10 to 98-362-12, 2002WL 27305 (E. D. Pa., Januray 7, 2002), rev'd Nos. 98-362-10, 98-362-11, 98-362-12, 2002 WL 389163 (E. D. Pa. March 13, 2002)

² Moran, B. and Murdock, J., Joseph Ramirez vs. State of Florida – Supreme Court Decision, December 20, 1991, AFTE Journal, 2002: 34(2): 215 – 226

³ Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993)

⁴ Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999)

⁵ Grzybowski, R. and Murdock, J., Firearm and Toolmark Identification – Meeting the Daubert Challenge. AFTE Journal 1998: 30(1): 3-14.

⁶ Frye v. United States, 293 F. 1013 (D.C. Cir. 1923)

⁷ Gunther, C.O., “Markings on Bullets and Shells Fired from Small Arms”, Mechanical Engineering #54, 1932, 341 –345 – at 334

⁸ Hatcher, J.S., Textbook of Firearms Investigation, Identification and Evidence. Marines, NC: Small Arms Technical Publishing Company, 1935, 286.

⁹ Nichols, R., Firearm and Toolmark Identification Criteria: A Review of the Literature, Journal of Forensic Sciences 1997 May; 42(3): 466-474.

¹⁰ Nichols, R., Firearm and Toolmark Identification Criteria: A Review of the Literature, Part II, Journal of Forensic Sciences 2003 March; 48(2):318-327

¹¹“Theory of Identification, Range of Striae Comparison Reports and Modified Glossary Definitions –An AFTE Criteria for Identification Committee Report, AFTE Journal 1992, 24(3): 337

¹² Ibid, pp 337 -338

¹³ Walsh, K., and Wevers, G., – Toolmark Identification: Can We Determine a Criterion?, Science and Justice: JFSS Interfaces, Jan – March 2002, (29): 4 – 5

¹⁴ Stone, R., “How Unique Are Impressed Toolmarks?, submitted for publication in the AFTE Journal, February 2003.

¹⁵ Biasotti, AA, Murdock J., Firearm and Toolmark Identification, Chapter 23, Vol. 2, pp 140 in Modern Scientific Evidence: The Law and Science of Expert Testimony, (Faigman DL, Kaye DK, Saks MJ, Sanders, J. eds., 1997). St. Paul: West, (unchanged on pp. 516 of 2002, 2nd Edition)

¹⁶ Souder, W., Firearms Identification, AFTE Journal 1985, 17(3): 114 – 118, (originally published in the Army and Navy Journal, March 19, 1932, pp 675, 677)

¹⁷ Thornton, J., “The Validity of Firearms Evidence”, California Attorneys for Criminal Justice Forum, Vol. 5, No. 4, p 15 – 25, July-Aug 1978 and AFTE Journal, Vol. 11 No. 2, April 1979, p 16 – 19

¹⁸ Peterson, J., and Markham, P., Crime laboratory testing results, 1978 – 1991, II: Resolving Questions of Common Origin. J of Forensic Sci 1995: 40(6): 1009 – 1029.

¹⁹ Grzybowski, R., and Murdock, J., op.cit.

²⁰ Ibid

²¹ Ibid

²² Biasotti, A., and Murdock, J., Firearm and Toolmark Identification, Chapter 29, Vol. 3 in *Modern Scientific Evidence: The Law and Science of Expert Testimony* (Faigman DL, Kaye DK, Saks MJ, Sanders, J., eds., 2nd ed, 2002), St. Paul: West

²³ Ibid

²⁴ Moenssens, A., Meeting the Daubert Challenge to Handwriting Evidence – Preparing for a Daubert Hearing, abstract of a talk given at the Second Annual Symposium on the Forensic Examination of Questioned Documents at Albany, N.Y. on June 18, 1999. (see www.forensic-evidence.com) An earlier version of this also appears in the October, 1999 issue of the *Forensic Science Communications*, a peer-reviewed quarterly journal published on the Internet by FBI Laboratory personnel (See <http://www.fbi.gov/programs/lab/fsc/current/index.htm>)

²⁵ Denio, D., The History of the AFTE Journal, the Peer Review Process, and Daubert Issues, AFTE Journal 2002, 34(2):210-214

²⁶ Buxton, J., The Science of Ballistics: Judicial Applications, American Journal of Police Science, May – June, 1931, 2(3): 211 - 219

²⁷ Serhant, J., The Admissibility of Ballistics in Evidence, American Journal of Police Science, May – June, 1931, 2(3): 202 - 210

²⁸ Baker, N., The Campbell Case, American Journal of Police Science, Jan – Feb, 1932, 3(1): 21 - 31

²⁹ Inbau, F., Scientific Evidence in Criminal Cases (Firearms Identification – “Ballistics”), AFTE Journal 1981, 13(2): 281 (Originally appeared in Journal of Police Science, 1933)

³⁰ Gunther, A.B, and Gunther, M.E., The Identification of Firearms, John Wiley and Sons, Inc, London: Champman and Hall, Limited, 1935, see chapter 3 – p. 246 – 327, “Analysis of the authorities on Firearms Identification”

³¹ Goddard, C., A History of Firearm Identification, AFTE Journal, 1985, 17(1): 55 – 68 (originally printed in Chicago Police Journal, 1936)

³² FBI, Firearms Identification, U.S. Government Printing Office, 1941, see pp 17 – 33 “Development and Admissibility of Ballistics and Firearms Evidence”

³³ Goddard, C., A History of Firearms Identification to 1930, AFTE Journal 1999, 31(3): 225 – 241

³⁴ IAI Firearms Sub-Committee –Stanton O. Berg, Chairman, The History of Firearms Identification, Identification News, June 1965, pp 5 - 15

³⁵ Saferstein, R., Forensic Science Handbook- Volume II, Prentice Hall, 1988, see p 411- 416 – “Firearms Identification – Historical Background”

³⁶ Hamby, J., The History of Firearm and Toolmark Identification, AFTE Journal 1999, 31(3): 266 – 284

³⁷ Kraft, B., Critical Review of Forensic Ballistics – Part I, American Journal of Police Science, Jan– Feb 1931, 2(1): 52 - 66

³⁸ Kraft, B., Critical Review of Forensic Ballistics – Part II, American Journal of Police Science, Mar – Apr, 1931, 2(2): 125 - 142